

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
25 May 2001 (25.05.2001)

PCT

(10) International Publication Number  
**WO 01/36444 A1**

(51) International Patent Classification<sup>7</sup>: **C07H 21/04**,  
C12N 5/10, 15/29, 15/63, 15/82

[CN/US]; 242 Gravatt Drive, Berkeley, CA 94705 (US).  
**JIANG, Cai-Zhong** [CN/US]; 34495 Heathrow Terrace,  
Fremont, CA 94555 (US).

(21) International Application Number: PCT/US00/31325

(74) Common Representative: **MENDEL BIOTECHNOLOGY, INC.**; Guerrero, Karen, 21375 Cabot Boulevard,  
Hayward, CA 94545 (US).

(22) International Filing Date:  
14 November 2000 (14.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/166,228 17 November 1999 (17.11.1999) US  
60/197,899 17 April 2000 (17.04.2000) US  
60/227,439 22 August 2000 (22.08.2000) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*):  
**MENDEL BIOTECHNOLOGY, INC.** [US/US]; 21375  
Cabot Boulevard, Hayward, CA 94541 (US).

(71) Applicants and

(72) Inventors: **RIECHMANN, Jose Luis** [ES/US]; 115 Moss  
Avenue #308, Oakland, CA 94611 (US). **REUBER, Lynne**  
[US/US]; 2000 Walnut Avenue, Fremont, CA 94538 (US).  
**KEDDIE, James** [GB/US]; 54 McLellan Avenue, San  
Mateo, CA 94403 (US). **RATCLIFFE, Oliver** [GB/US];  
814 East 21st Street, Oakland, CA 94606 (US). **HEARD,**  
**Jacqueline** [US/US]; 810 Guildford Avenue, San Mateo,  
CA 94402 (US). **SAMAH, Raymond** [US/US]; 2224  
Albert Lane, Capitola, CA 95010 (US). **YU, Guo-Liang**

Published:

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PLANT DEVELOPMENTAL GENES

(57) Abstract: Recombinant polynucleotides and methods for modifying the phenotype of a plant are provided. In particular, the phenotype that is being modified is a plant's structure and development characteristics.



WO 01/36444 A1

<b>PLANT DEVELOPMENTAL GENES</b>
----------------------------------

**RELATED APPLICATION INFORMATION**

5 The present invention claims the benefit from US Provisional Patent Application Serial Nos. 60/166,228 filed November 17, 1999 and 60/197,899 filed April 17, 2000 and "Plant Trait Modification III" filed August 22, 2000.

**FIELD OF THE INVENTION**

This invention relates to the field of plant biology. More particularly, the present invention pertains to compositions and methods for phenotypically modifying a plant.

10

**BACKGROUND OF THE INVENTION**

Transcription factors can modulate gene expression, either increasing or decreasing (inducing or repressing) the rate of transcription. This modulation results in differential levels of gene expression at various developmental stages, in different tissues and cell types, and in response to different exogenous (e.g., environmental) and endogenous stimuli throughout the life cycle of the organism.

15

Because transcription factors are key controlling elements of biological pathways, altering the expression levels of one or more transcription factors can change entire biological pathways in an organism. For example, manipulation of the levels of selected transcription factors may result in increased expression of economically useful proteins or metabolic chemicals in plants or to improve other agriculturally relevant characteristics. Conversely, blocked or reduced expression of a transcription factor may reduce biosynthesis of unwanted compounds or remove an undesirable trait. Therefore, manipulating transcription factor levels in a plant offers tremendous potential in agricultural biotechnology for modifying a plant's traits.

20

25

The present invention provides novel transcription factors useful for modifying a plant's phenotype in desirable ways, such as modifying a plant's structure or development.

**SUMMARY OF THE INVENTION**

In a first aspect, the invention relates to a recombinant polynucleotide comprising a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a polypeptide comprising a sequence selected from SEQ ID Nos. 2N, where N=1-23, or a complementary nucleotide sequence thereof; (b) a nucleotide sequence encoding a polypeptide comprising a conservatively substituted variant of a polypeptide of (a); (c) a nucleotide sequence comprising a sequence selected from those of SEQ ID Nos. 2N-1, where N=1-23, or a

30

complementary nucleotide sequence thereof; (d) a nucleotide sequence comprising silent substitutions in a nucleotide sequence of (c); (e) a nucleotide sequence which hybridizes under stringent conditions over substantially the entire length of a nucleotide sequence of one or more of: (a), (b), (c), or (d); (f) a nucleotide sequence comprising at least 15 consecutive nucleotides of  
5 a sequence of any of (a)-(e); (g) a nucleotide sequence comprising a subsequence or fragment of any of (a)-(f), which subsequence or fragment encodes a polypeptide having a biological activity that modifies a plant's structure and development characteristics; (h) a nucleotide sequence having at least 31% sequence identity to a nucleotide sequence of any of (a)-(g); (i) a nucleotide sequence having at least 60% identity sequence identity to a nucleotide sequence of any of (a)-  
10 (g); (j) a nucleotide sequence which encodes a polypeptide having at least 31% identity sequence identity to a polypeptide of SEQ ID Nos. 2N, where N=1-23; (k) a nucleotide sequence which encodes a polypeptide having at least 60% identity sequence identity to a polypeptide of SEQ ID Nos. 2N, where N=1-23; and (l) a nucleotide sequence which encodes a conserved domain of a polypeptide having at least 65% sequence identity to a conserved domain of a polypeptide of  
15 SEQ ID Nos. 2N, where N=1-23. The recombinant polynucleotide may further comprise a constitutive, inducible, or tissue-active promoter operably linked to the nucleotide sequence. The invention also relates to compositions comprising at least two of the above described polynucleotides.

In a second aspect, the invention is an isolated or recombinant polypeptide  
20 comprising a subsequence of at least about 15 contiguous amino acids encoded by the recombinant or isolated polynucleotide described above.

In another aspect, the invention is a transgenic plant comprising one or more of the above described recombinant polynucleotides. In yet another aspect, the invention is a plant with altered expression levels of a polynucleotide described above or a plant with altered  
25 expression or activity levels of an above described polypeptide. Further, the invention is a plant lacking a nucleotide sequence encoding a polypeptide described above. The plant may be a soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honeydew, lettuce, mango, melon, onion, papaya, peas, peppers,  
30 pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits, or vegetable brassicas plant.

In a further aspect, the invention relates to a cloning or expression vector comprising the isolated or recombinant polynucleotide described above or cells comprising the cloning or expression vector.

In yet a further aspect, the invention relates to a composition produced by incubating a polynucleotide of the invention with a nuclease, a restriction enzyme, a polymerase; a polymerase and a primer; a cloning vector, or with a cell.

5 Furthermore, the invention relates to a method for producing a plant having modified structure and development traits. The method comprises altering the expression of an isolated or recombinant polynucleotide of the invention or altering the expression or activity of a polypeptide of the invention in a plant to produce a modified plant, and selecting the modified plant for modified structure and development traits.

10 In another aspect, the invention relates to a method of identifying a factor that is modulated by or interacts with a polypeptide encoded by a polynucleotide of the invention. The method comprises expressing a polypeptide encoded by the polynucleotide in a plant; and identifying at least one factor that is modulated by or interacts with the polypeptide. In one embodiment the method for identifying modulating or interacting factors is by detecting binding by the polypeptide to a promoter sequence, or by detecting interactions between an additional  
15 protein and the polypeptide in a yeast two hybrid system, or by detecting expression of a factor by hybridization to a microarray, subtractive hybridization or differential display.

In yet another aspect, the invention is a method of identifying a molecule that modulates activity or expression of a polynucleotide or polypeptide of interest. The method comprises placing the molecule in contact with a plant comprising the polynucleotide or  
20 polypeptide encoded by the polynucleotide of the invention and monitoring one or more of the expression level of the polynucleotide in the plant, the expression level of the polypeptide in the plant, and modulation of an activity of the polypeptide in the plant.

In yet another aspect, the invention relates to an integrated system, computer or computer readable medium comprising one or more character strings corresponding to a  
25 polynucleotide of the invention, or to a polypeptide encoded by the polynucleotide. The integrated system, computer or computer readable medium may comprise a link between one or more sequence strings to a modified plant structure and development trait.

In yet another aspect, the invention is a method for identifying a sequence similar or homologous to one or more polynucleotides of the invention, or one or more polypeptides  
30 encoded by the polynucleotides. The method comprises providing a sequence database; and, querying the sequence database with one or more target sequences corresponding to the one or more polynucleotides or to the one or more polypeptides to identify one or more sequence members of the database that display sequence similarity or homology to one or more of the one or more target sequences.



The method may further comprise of linking the one or more of the polynucleotides of the invention, or encoded polypeptides, to a modified plant structure and development characteristics phenotype.

#### BRIEF DESCRIPTION OF THE DRAWINGS

5        Figure 1 provides a table of exemplary polynucleotide and polypeptide sequences of the invention. The table includes from left to right for each sequence: the SEQ ID No., the internal code reference number (GID), whether the sequence is a polynucleotide or polypeptide sequence, and identification of any conserved domains for the polypeptide sequences.

10        Figure 2 provides a table of exemplary sequences that are homologous to other sequences provided in the Sequence Listing and that are derived from *Arabidopsis thaliana*. The table includes from left to right: the SEQ ID No., the internal code reference number (GID), identification of the homologous sequence, whether the sequence is a polynucleotide or polypeptide sequence, and identification of any conserved domains for the polypeptide sequences.

15        Figure 3 provides a table of exemplary sequences that are homologous to the sequences provided in Figures 1 and 2 and that are derived from plants other than *Arabidopsis thaliana*. The table includes from left to right: the SEQ ID No., the internal code reference number (GID), the unique GenBank sequence ID No. (NID), the probability that the comparison was generated by chance (P-value), and the species from which the homologous gene was identified.

20

#### DETAILED DESCRIPTION

The present invention relates to polynucleotides and polypeptides, e.g. for modifying phenotypes of plants.

25        In particular, the polynucleotides or polypeptides are useful for modifying traits associated with a plant's structure or development characteristics when the expression levels of the polynucleotides or expression levels or activity levels of the polypeptides are altered. Specifically, the polynucleotides and polypeptides are useful for modifying the structure and size of flowers, leaves, roots, the plant as a whole, or the like, apical dominance, branching patterns, number of organs, organ identity, whether a plant is sterile or not, the vascularization of a plant, 30 or the developmental staging of a plant, such as when senescence is triggered.

The polynucleotides of the invention encode plant transcription factors. The plant transcription factors are derived, e.g., from *Arabidopsis thaliana* and can belong, e.g., to one or more of the following transcription factor families: the AP2 (APETALA2) domain transcription

factor family (Riechmann and Meyerowitz (1998) J. Biol. Chem. 379:633-646); the MYB transcription factor family (Martin and Paz-Ares (1997) Trends Genet. 13:67-73); the MADS domain transcription factor family (Riechmann and Meyerowitz (1997) J. Biol. Chem. 378:1079-1101); the WRKY protein family (Ishiguro and Nakamura (1994) Mol. Gen. Genet. 244:563-571); the ankyrin-repeat protein family (Zhang et al. (1992) Plant Cell 4:1575-1588); the miscellaneous protein (MISC) family (Kim et al. (1997) Plant J. 11:1237-1251); the zinc finger protein (Z) family (Klug and Schwabe (1995) FASEB J. 9: 597-604); the homeobox (HB) protein family (Duboule (1994) Guidebook to the Homeobox Genes, Oxford University Press); the CAAT-element binding proteins (Forsburg and Guarente (1989) Genes Dev. 3:1166-1178); the squamosa promoter binding proteins (SPB) (Klein et al. (1996) Mol. Gen. Genet. 1996 250:7-16); the NAM protein family; the IAA/AUX proteins (Rouse et al. (1998) Science 279:1371-1373); the HLH/MYC protein family (Littlewood et al. (1994) Prot. Profile 1:639-709); the DNA-binding protein (DBP) family (Tucker et al. (1994) EMBO J. 13:2994-3002); the bZIP family of transcription factors (Foster et al. (1994) FASEB J. 8:192-200); the BPF-1 protein (Box P-binding factor) family (da Costa e Silva et al. (1993) Plant J. 4:125-135); and the golden protein (GLD) family (Hall et al. (1998) Plant Cell 10:925-936).

In addition to methods for modifying a plant phenotype by employing one or more polynucleotides and polypeptides of the invention described herein, the polynucleotides and polypeptides of the invention have a variety of additional uses. These uses include their use in the recombinant production (i.e., expression) of proteins; as regulators of plant gene expression, as diagnostic probes for the presence of complementary or partially complementary nucleic acids (including for detection of natural coding nucleic acids); as substrates for further reactions, e.g., mutation reactions, PCR reactions, or the like, or as substrates for cloning e.g., including digestion or ligation reactions, and for identifying exogenous or endogenous modulators of the transcription factors.

## DEFINITIONS

A "polynucleotide" is a nucleic acid sequence comprising a plurality of polymerized nucleotide residues, e.g., at least about 15 consecutive polymerized nucleotide residues, optionally at least about 30 consecutive nucleotides, at least about 50 consecutive nucleotides. In many instances, a polynucleotide comprises a nucleotide sequence encoding a polypeptide (or protein) or a domain or fragment thereof. Additionally, the polynucleotide may comprise a promoter, an intron, an enhancer region, a polyadenylation site, a translation initiation site, 5' or 3' untranslated regions, a reporter gene, a selectable marker, or the like. The

polynucleotide can be single stranded or double stranded DNA or RNA. The polynucleotide optionally comprises modified bases or a modified backbone. The polynucleotide can be, e.g., genomic DNA or RNA, a transcript (such as an mRNA), a cDNA, a PCR product, a cloned DNA, a synthetic DNA or RNA, or the like. The polynucleotide can comprise a sequence in either  
5 sense or antisense orientations.

A "recombinant polynucleotide" is a polynucleotide that is not in its native state, e.g., the polynucleotide comprises a nucleotide sequence not found in nature, or the polynucleotide is in a context other than that in which it is naturally found, e.g., separated from nucleotide sequences with which it typically is in proximity in nature, or adjacent (or contiguous  
10 with) nucleotide sequences with which it typically is not in proximity. For example, the sequence at issue can be cloned into a vector, or otherwise recombined with one or more additional nucleic acid.

An "isolated polynucleotide" is a polynucleotide whether naturally occurring or recombinant, that is present outside the cell in which it is typically found in nature, whether  
15 purified or not. Optionally, an isolated polynucleotide is subject to one or more enrichment or purification procedures, e.g., cell lysis, extraction, centrifugation, precipitation, or the like.

A "recombinant polypeptide" is a polypeptide produced by translation of a recombinant polynucleotide. An "isolated polypeptide," whether a naturally occurring or a recombinant polypeptide, is more enriched in (or out of) a cell than the polypeptide in its natural  
20 state in a wild type cell, e.g., more than about 5% enriched, more than about 10% enriched, or more than about 20%, or more than about 50%, or more, enriched, i.e., alternatively denoted: 105%, 110%, 120%, 150% or more, enriched relative to wild type standardized at 100%. Such an enrichment is not the result of a natural response of a wild type plant. Alternatively, or additionally, the isolated polypeptide is separated from other cellular components with which it is  
25 typically associated, e.g., by any of the various protein purification methods herein.

The term "transgenic plant" refers to a plant that contains genetic material, not found in a wild type plant of the same species, variety or cultivar. The genetic material may include a transgene, an insertional mutagenesis event (such as by transposon or T-DNA insertional mutagenesis), an activation tagging sequence, a mutated sequence, a homologous  
30 recombination event or a sequence modified by chimeraplasty. Typically, the foreign genetic material has been introduced into the plant by human manipulation.

A transgenic plant may contain an expression vector or cassette. The expression cassette typically comprises a polypeptide-encoding sequence operably linked (i.e., under regulatory control of) to appropriate inducible or constitutive regulatory sequences that allow for

the expression of polypeptide. The expression cassette can be introduced into a plant by transformation or by breeding after transformation of a parent plant. A plant refers to a whole plant as well as to a plant part, such as seed, fruit, leaf, or root, plant tissue, plant cells or any other plant material, e.g., a plant explant, as well as to progeny thereof, and to *in vitro* systems that mimic biochemical or cellular components or processes in a cell.

The phrase "ectopically expression or altered expression" in reference to a polynucleotide indicates that the pattern of expression in, e.g., a transgenic plant or plant tissue, is different from the expression pattern in a wild type plant or a reference plant of the same species. For example, the polynucleotide or polypeptide is expressed in a cell or tissue type other than a cell or tissue type in which the sequence is expressed in the wild type plant, or by expression at a time other than at the time the sequence is expressed in the wild type plant, or by a response to different inducible agents, such as hormones or environmental signals, or at different expression levels (either higher or lower) compared with those found in a wild type plant. The term also refers to altered expression patterns that are produced by lowering the levels of expression to below the detection level or completely abolishing expression. The resulting expression pattern can be transient or stable, constitutive or inducible. In reference to a polypeptide, the term "ectopic expression or altered expression" further may relate to altered activity levels resulting from the interactions of the polypeptides with exogenous or endogenous modulators or from interactions with factors or as a result of the chemical modification of the polypeptides.

The term "fragment" or "domain," with respect to a polypeptide, refers to a subsequence of the polypeptide. In some cases, the fragment or domain, is a subsequence of the polypeptide which performs at least one biological function of the intact polypeptide in substantially the same manner, or to a similar extent, as does the intact polypeptide. For example, a polypeptide fragment can comprise a recognizable structural motif or functional domain such as a DNA binding domain that binds to a DNA promoter region, an activation domain or a domain for protein-protein interactions. Fragments can vary in size from as few as 6 amino acids to the full length of the intact polypeptide, but are preferably at least about 30 amino acids in length and more preferably at least about 60 amino acids in length. In reference to a nucleotide sequence, "a fragment" refers to any subsequence of a polynucleotide, typically, of at least consecutive about 15 nucleotides, preferably at least about 30 nucleotides, more preferably at least about 50, of any of the sequences provided herein.

The term "trait" refers to a physiological, morphological, biochemical or physical characteristic of a plant or particular plant material or cell. In some instances, this characteristic is visible to the human eye, such as seed or plant size, or can be measured by available

biochemical techniques, such as the protein, starch or oil content of seed or leaves or by the observation of the expression level of genes, e.g., by employing Northern analysis, RT-PCR, microarray gene expression assays or reporter gene expression systems, or by agricultural observations such as stress tolerance, yield or pathogen tolerance.

5                   “Trait modification” refers to a detectable difference in a characteristic in a plant ectopically expressing a polynucleotide or polypeptide of the present invention relative to a plant not doing so, such as a wild type plant. In some cases, the trait modification can be evaluated quantitatively. For example, the trait modification can entail at least about a 2% increase or decrease in an observed trait (difference), at least a 5% difference, at least about a 10%  
10 difference, at least about a 20% difference, at least about a 30%, at least about a 50%, at least about a 70%, or at least about a 100%, or an even greater difference. It is known that there can be a natural variation in the modified trait. Therefore, the trait modification observed entails a change of the normal distribution of the trait in the plants compared with the distribution observed in wild type plant.

15                   Trait modifications of particular interest include those to seed (such as embryo or endosperm), fruit, root, flower, leaf, stem, shoot, seedling or the like, including: enhanced tolerance to environmental conditions including freezing, chilling, heat, drought, water saturation, radiation and ozone; improved tolerance to microbial, fungal or viral diseases; improved tolerance to pest infestations, including nematodes, mollicutes, parasitic higher plants or the like;  
20 decreased herbicide sensitivity; improved tolerance of heavy metals or enhanced ability to take up heavy metals; improved growth under poor photoconditions (e.g., low light and/or short day length), or changes in expression levels of genes of interest. Other phenotype that can be modified relate to the production of plant metabolites, such as variations in the production of taxol, tocopherol, tocotrienol, sterols, phytosterols, vitamins, wax monomers, anti-oxidants,  
25 amino acids, lignins, cellulose, tannins, prenillipids (such as chlorophylls and carotenoids), glucosinolates, and terpenoids, enhanced or compositionally altered protein or oil production (especially in seeds), or modified sugar (insoluble or soluble) and/or starch composition. Physical plant characteristics that can be modified include cell development (such as the number of trichomes), fruit and seed size and number, yields of plant parts such as stems, leaves and  
30 roots, the stability of the seeds during storage, characteristics of the seed pod (e.g., susceptibility to shattering), root hair length and quantity, internode distances, or the quality of seed coat. Plant growth characteristics that can be modified include growth rate, germination rate of seeds, vigor of plants and seedlings, leaf and flower senescence, male sterility, apomixis, flowering time, flower abscission, rate of nitrogen uptake, biomass or transpiration characteristics, as well as

plant architecture characteristics such as apical dominance, branching patterns, number of organs, organ identity, organ shape or size.

#### POLYPEPTIDES AND POLYNUCLEOTIDES OF THE INVENTION

5 The present invention provides, among other things, transcription factors (TFs), and transcription factor homologue polypeptides, and isolated or recombinant polynucleotides encoding the polypeptides. These polypeptides and polynucleotides may be employed to modify a plant's structure and development characteristics.

10 Exemplary polynucleotides encoding the polypeptides of the invention were identified in the *Arabidopsis thaliana* GenBank database using publicly available sequence analysis programs and parameters. Sequences initially identified were then further characterized to identify sequences comprising specified sequence strings corresponding to sequence motifs present in families of known transcription factors. Polynucleotide sequences meeting such criteria were confirmed as transcription factors.

15 Additional polynucleotides of the invention were identified by screening *Arabidopsis thaliana* and/or other plant cDNA libraries with probes corresponding to known transcription factors under low stringency hybridization conditions. Additional sequences, including full length coding sequences were subsequently recovered by the rapid amplification of cDNA ends (RACE) procedure, using a commercially available kit according to the manufacturer's instructions. Where necessary, multiple rounds of RACE are performed to isolate 5' and 3' ends. The full length cDNA was then recovered by a routine end-to-end polymerase chain reaction (PCR) using primers specific to the isolated 5' and 3' ends. Exemplary sequences are provided in the Sequence Listing.

20 The polynucleotides of the invention were ectopically expressed in overexpressor or knockout plants and changes in the structure and development characteristics of the plants were observed. Therefore, the polynucleotides and polypeptides can be employed to improve the structure and development characteristics of plants.

#### Making polynucleotides

30 The polynucleotides of the invention include sequences that encode transcription factors and transcription factor homologue polypeptides and sequences complementary thereto, as well as unique fragments of coding sequence, or sequence complementary thereto. Such polynucleotides can be, e.g., DNA or RNA, e.g., mRNA, cRNA, synthetic RNA, genomic DNA, cDNA synthetic DNA, oligonucleotides, etc. The polynucleotides are either double-stranded or single-stranded, and include either, or both sense (i.e., coding) sequences and antisense (i.e., non-

coding, complementary) sequences. The polynucleotides include the coding sequence of a transcription factor, or transcription factor homologue polypeptide, in isolation, in combination with additional coding sequences (e.g., a purification tag, a localization signal, as a fusion-protein, as a pre-protein, or the like), in combination with non-coding sequences (e.g., introns or  
5     intons, regulatory elements such as promoters, enhancers, terminators, and the like), and/or in a vector or host environment in which the polynucleotide encoding a transcription factor or transcription factor homologue polypeptide is an endogenous or exogenous gene.

A variety of methods exist for producing the polynucleotides of the invention. Procedures for identifying and isolating DNA clones are well known to those of skill in the art, and are described in, e.g., Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods  
10     in Enzymology volume 152 Academic Press, Inc., San Diego, CA ("Berger"); Sambrook et al., Molecular Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook") and Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing  
15     Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2000) ("Ausubel").

Alternatively, polynucleotides of the invention, can be produced by a variety of in vitro amplification methods adapted to the present invention by appropriate selection of specific or degenerate primers. Examples of protocols sufficient to direct persons of skill through in vitro amplification methods, including the polymerase chain reaction (PCR) the ligase chain  
20     reaction (LCR), Qbeta-replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA), e.g., for the production of the homologous nucleic acids of the invention are found in Berger, Sambrook, and Ausubel, as well as Mullis et al., (1987) PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press Inc. San Diego, CA (1990) (Innis). Improved methods for cloning in vitro amplified nucleic acids are described in Wallace et al.,  
25     U.S. Pat. No. 5,426,039. Improved methods for amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369: 684-685 and the references cited therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. See, e.g., Ausubel,  
30     Sambrook and Berger, *all supra*.

Alternatively, polynucleotides and oligonucleotides of the invention can be assembled from fragments produced by solid-phase synthesis methods. Typically, fragments of up to approximately 100 bases are individually synthesized and then enzymatically or chemically ligated to produce a desired sequence, e.g., a polynucleotide encoding all or part of a

transcription factor. For example, chemical synthesis using the phosphoramidite method is described, e.g., by Beaucage et al. (1981) Tetrahedron Letters 22:1859-69; and Matthes et al. (1984) EMBO J. 3:801-5. According to such methods, oligonucleotides are synthesized, purified, annealed to their complementary strand, ligated and then optionally cloned into suitable vectors.

- 5 And if so desired, the polynucleotides and polypeptides of the invention can be custom ordered from any of a number of commercial suppliers.

#### HOMOLOGOUS SEQUENCES

- Sequences homologous, i.e., that share significant sequence identity or similarity, to those provided in the Sequence Listing, derived from *Arabidopsis thaliana* or from other plants of choice are also an aspect of the invention. Homologous sequences can be derived from any plant including monocots and dicots and in particular agriculturally important plant species, including but not limited to, crops such as soybean, wheat, corn, potato, cotton, rice, oilseed rape (including canola), sunflower, alfalfa, sugarcane and turf; or fruits and vegetables, such as banana, blackberry, blueberry, strawberry, and raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honeydew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits (such as apple, peach, pear, cherry and plum) and vegetable brassicas (such as broccoli, cabbage, cauliflower, brussel sprouts and kohlrabi). Other crops, fruits and vegetables whose phenotype can be changed include barley, rye, millet, sorghum, currant, avocado, citrus fruits such as oranges, lemons, grapefruit and tangerines, artichoke, cherries, nuts such as the walnut and peanut, endive, leek, roots, such as arrowroot, beet, cassava, turnip, radish, yam, and sweet potato, and beans. The homologous sequences may also be derived from woody species, such as pine, poplar and eucalyptus.

- Transcription factors that are homologous to the listed sequences will typically share at least about 30% amino acid sequence identity. More closely related transcription factors can share at least about 50%, about 60%, about 65%, about 70%, about 75% or about 80% or about 90% or about 95% or about 98% or more sequence identity with the listed sequences. Factors that are most closely related to the listed sequences share, e.g., at least about 85%, about 90% or about 95% or more % sequence identity to the listed sequences. At the nucleotide level, the sequences will typically share at least about 40% nucleotide sequence identity, preferably at least about 50%, about 60%, about 70% or about 80% sequence identity, and more preferably about 85%, about 90%, about 95% or about 97% or more sequence identity to one or more of the listed sequences. The degeneracy of the genetic code enables major variations in the nucleotide



sequence of a polynucleotide while maintaining the amino acid sequence of the encoded protein. Conserved domains within a transcription factor family may exhibit a higher degree of sequence homology, such as at least 65% sequence identity including conservative substitutions, and preferably at least 80% sequence identity.

5                    Identifying Nucleic Acids by Hybridization

Polynucleotides homologous to the sequences illustrated in the Sequence Listing can be identified, e.g., by hybridization to each other under stringent or under highly stringent conditions. Single stranded polynucleotides hybridize when they associate based on a variety of well characterized physico-chemical forces, such as hydrogen bonding, solvent exclusion, base  
10                    stacking and the like. The stringency of a hybridization reflects the degree of sequence identity of the nucleic acids involved, such that the higher the stringency, the more similar are the two polynucleotide strands. Stringency is influenced by a variety of factors, including temperature, salt concentration and composition, organic and non-organic additives, solvents, etc. present in both the hybridization and wash solutions and incubations (and number), as described in more  
15                    detail in the references cited above.

An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined  
20                    ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Nucleic acid molecules that hybridize under stringent conditions will typically hybridize to a probe based on either the entire cDNA or selected portions, e.g., to a unique subsequence, of the cDNA under wash conditions of 0.2x SSC to 2.0 x SSC, 0.1% SDS at 50-65° C, for example 0.2 x SSC, 0.1% SDS at 65° C. For identification of less closely related homologues washes can  
25                    be performed at a lower temperature, e.g., 50° C. In general, stringency is increased by raising the wash temperature and/or decreasing the concentration of SSC.

As another example, stringent conditions can be selected such that an oligonucleotide that is perfectly complementary to the coding oligonucleotide hybridizes to the coding oligonucleotide with at least about a 5-10x higher signal to noise ratio than the ratio for  
30                    hybridization of the perfectly complementary oligonucleotide to a nucleic acid encoding a transcription factor known as of the filing date of the application. Conditions can be selected such that a higher signal to noise ratio is observed in the particular assay which is used, e.g., about 15x, 25x, 35x, 50x or more. Accordingly, the subject nucleic acid hybridizes to the unique coding oligonucleotide with at least a 2x higher signal to noise ratio as compared to hybridization

of the coding oligonucleotide to a nucleic acid encoding known polypeptide. Again, higher signal to noise ratios can be selected, e.g., about 5x, 10x, 25x, 35x, 50x or more. The particular signal will depend on the label used in the relevant assay, e.g., a fluorescent label, a colorimetric label, a radio active label, or the like.

5                   Alternatively, transcription factor homologue polypeptides can be obtained by screening an expression library using antibodies specific for one or more transcription factors. With the provision herein of the disclosed transcription factor, and transcription factor homologue nucleic acid sequences, the encoded polypeptide(s) can be expressed and purified in a heterologous expression system (e.g., *E. coli*) and used to raise antibodies (monoclonal or  
10 polyclonal) specific for the polypeptide(s) in question. Antibodies can also be raised against synthetic peptides derived from transcription factor, or transcription factor homologue, amino acid sequences. Methods of raising antibodies are well known in the art and are described in Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, New York. Such antibodies can then be used to screen an expression library produced from the plant  
15 from which it is desired to clone additional transcription factor homologues, using the methods described above. The selected cDNAs can be confirmed by sequencing and enzymatic activity.

#### SEQUENCE VARIATIONS

It will readily be appreciated by those of skill in the art, that any of a variety of polynucleotide sequences are capable of encoding the transcription factors and transcription  
20 factor homologue polypeptides of the invention. Due to the degeneracy of the genetic code, many different polynucleotides can encode identical and/or substantially similar polypeptides in addition to those sequences illustrated in the Sequence Listing.

For example, Table 1 illustrates, e.g., that the codons AGC, AGT, TCA, TCC, TCG, and TCT all encode the same amino acid: serine. Accordingly, at each position in the  
25 sequence where there is a codon encoding serine, any of the above trinucleotide sequences can be used without altering the encoded polypeptide.

**Table 1**

Amino acids			Codon							
Alanine	Ala	A	GCA	GCC	GCG	GCU				
Cysteine	Cys	C	TGC	TGT						
Aspartic acid	Asp	D	GAC	GAT						
Glutamic acid	Glu	E	GAA	GAG						
Phenylalanine	Phe	F	TTC	TTT						
Glycine	Gly	G	GGA	GGC	GGG	GGT				
Histidine	His	H	CAC	CAT						
Isoleucine	Ile	I	ATA	ATC	ATT					
Lysine	Lys	K	AAA	AAG						
Leucine	Leu	L	TTA	TTG	CTA	CTC	CTG	CTT		
Methionine	Met	M	ATG							
Asparagine	Asn	N	AAC	AAT						
Proline	Pro	P	CCA	CCC	CCG	CCT				
Glutamine	Gln	Q	CAA	CAG						
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGT		
Serine	Ser	S	AGC	AGT	TCA	TCC	TCG	TCT		
Threonine	Thr	T	ACA	ACC	ACG	ACT				
Valine	Val	V	GTA	GTC	GTG	GTT				
Tryptophan	Trp	W	TGG							
Tyrosine	Tyr	Y	TAC	TAT						

Sequence alterations that do not change the amino acid sequence encoded by the polynucleotide are termed "silent" variations. With the exception of the codons ATG and TGG, encoding methionine and tryptophan, respectively, any of the possible codons for the same amino acid can be substituted by a variety of techniques, e.g., site-directed mutagenesis, available in the art. Accordingly, any and all such variations of a sequence selected from the above table are a feature of the invention.

In addition to silent variations, other conservative variations that alter one, or a few amino acids in the encoded polypeptide, can be made without altering the function of the polypeptide, these conservative variants are, likewise, a feature of the invention.

For example, substitutions, deletions and insertions introduced into the sequences provided in the Sequence Listing are also envisioned by the invention. Such sequence modifications can be engineered into a sequence by site-directed mutagenesis (Wu (ed.) Meth. Enzymol. (1993) vol. 217, Academic Press) or the other methods noted below. Amino acid substitutions are typically of single residues; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. In preferred embodiments, deletions or insertions are made in adjacent pairs, e.g., a deletion of two residues or insertion of two residues. Substitutions, deletions, insertions or any combination thereof can be

combined to arrive at a sequence. The mutations that are made in the polynucleotide encoding the transcription factor should not place the sequence out of reading frame and should not create complementary regions that could produce secondary mRNA structure. Preferably, the polypeptide encoded by the DNA performs the desired function.

5                   Conservative substitutions are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the Table 2 when it is desired to maintain the activity of the protein. Table 2 shows amino acids which can be substituted for an amino acid in a protein and which are typically regarded as conservative substitutions.

10

**Table 2**

<b>Residue</b>	<b>Conservative Substitutions</b>
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Gln	Asn
Cys	Ser
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu, Val
Leu	Ile; Val
Lys	Arg; Gln
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr; Gly
Thr	Ser; Val
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Substitutions that are less conservative than those in Table 2 can be selected by picking residues that differ more significantly in their effect on maintaining (a) the structure of

the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in protein properties will be those in which (a) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine.

10 FURTHER MODIFYING SEQUENCES OF THE INVENTION—MUTATION/ FORCED EVOLUTION

In addition to generating silent or conservative substitutions as noted, above, the present invention optionally includes methods of modifying the sequences of the Sequence Listing. In the methods, nucleic acid or protein modification methods are used to alter the given sequences to produce new sequences and/or to chemically or enzymatically modify given sequences to change the properties of the nucleic acids or proteins.

Thus, in one embodiment, given nucleic acid sequences are modified, e.g., according to standard mutagenesis or artificial evolution methods to produce modified sequences. For example, Ausubel, *supra*, provides additional details on mutagenesis methods. Artificial forced evolution methods are described, e.g., by Stemmer (1994) *Nature* 370:389-391, and Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751. Many other mutation and evolution methods are also available and expected to be within the skill of the practitioner.

Similarly, chemical or enzymatic alteration of expressed nucleic acids and polypeptides can be performed by standard methods. For example, sequence can be modified by addition of lipids, sugars, peptides, organic or inorganic compounds, by the inclusion of modified nucleotides or amino acids, or the like. For example, protein modification techniques are illustrated in Ausubel, *supra*. Further details on chemical and enzymatic modifications can be found herein. These modification methods can be used to modify any given sequence, or to modify any sequence produced by the various mutation and artificial evolution modification methods noted herein.

Accordingly, the invention provides for modification of any given nucleic acid by mutation, evolution, chemical or enzymatic modification, or other available methods, as well as for the products produced by practicing such methods, e.g., using the sequences herein as a starting substrate for the various modification approaches.

For example, optimized coding sequence containing codons preferred by a particular prokaryotic or eukaryotic host can be used e.g., to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced using a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, preferred stop codons for *S. cerevisiae* and mammals are TAA and TGA, respectively. The preferred stop codon for monocotyledonous plants is TGA, whereas insects and *E. coli* prefer to use TAA as the stop codon.

The polynucleotide sequences of the present invention can also be engineered in order to alter a coding sequence for a variety of reasons, including but not limited to, alterations which modify the sequence to facilitate cloning, processing and/or expression of the gene product. For example, alterations are optionally introduced using techniques which are well known in the art, e.g., site-directed mutagenesis, to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to introduce splice sites, etc.

Furthermore, a fragment or domain derived from any of the polypeptides of the invention can be combined with domains derived from other transcription factors or synthetic domains to modify the biological activity of a transcription factor. For instance, a DNA binding domain derived from a transcription factor of the invention can be combined with the activation domain of another transcription factor or with a synthetic activation domain. A transcription activation domain assists in initiating transcription from a DNA binding site. Examples include the transcription activation region of VP16 or GAL4 (Moore et al. (1998) Proc. Natl. Acad. Sci. USA 95: 376-381; and Aoyama et al. (1995) Plant Cell 7:1773-1785), peptides derived from bacterial sequences (Ma and Ptashne (1987) Cell 51; 113-119) and synthetic peptides (Giniger and Ptashne, (1987) Nature 330:670-672).

## EXPRESSION AND MODIFICATION OF POLYPEPTIDES

Typically, polynucleotide sequences of the invention are incorporated into recombinant DNA (or RNA) molecules that direct expression of polypeptides of the invention in appropriate host cells, transgenic plants, in vitro translation systems, or the like. Due to the inherent degeneracy of the genetic code, nucleic acid sequences which encode substantially the same or a functionally equivalent amino acid sequence can be substituted for any listed sequence to provide for cloning and expressing the relevant homologue.

### Vectors, Promoters and Expression Systems

The present invention includes recombinant constructs comprising one or more of the nucleic acid sequences herein. The constructs typically comprise a vector, such as a plasmid, a cosmid, a phage, a virus (e.g., a plant virus), a bacterial artificial chromosome (BAC),  
5 a yeast artificial chromosome (YAC), or the like, into which a nucleic acid sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available.

10 General texts which describe molecular biological techniques useful herein, including the use and production of vectors, promoters and many other relevant topics, include Berger, Sambrook and Ausubel, *supra*. Any of the identified sequences can be incorporated into a cassette or vector, e.g., for expression in plants. A number of expression vectors suitable for stable transformation of plant cells or for the establishment of transgenic plants have been described  
15 including those described in Weissbach and Weissbach, (1989) Methods for Plant Molecular Biology, Academic Press, and Gelvin et al., (1990) Plant Molecular Biology Manual, Kluwer Academic Publishers. Specific examples include those derived from a Ti plasmid of *Agrobacterium tumefaciens*, as well as those disclosed by Herrera-Estrella et al. (1983) Nature 303: 209, Bevan (1984) Nucl Acid Res. 12: 8711-8721, Klee (1985) Bio/Technology 3: 637-642,  
20 for dicotyledonous plants.

Alternatively, non-Ti vectors can be used to transfer the DNA into monocotyledonous plants and cells by using free DNA delivery techniques. Such methods can involve, for example, the use of liposomes, electroporation, microprojectile bombardment, silicon  
carbide whiskers, and viruses. By using these methods transgenic plants such as wheat, rice  
25 (Christou (1991) Bio/Technology 9: 957-962) and corn (Gordon-Kamm (1990) Plant Cell 2: 603-618) can be produced. An immature embryo can also be a good target tissue for monocots for direct DNA delivery techniques by using the particle gun (Weeks et al. (1993) Plant Physiol 102: 1077-1084; Vasil (1993) Bio/Technology 10: 667-674; Wan and Lemeaux (1994) Plant Physiol 104: 37-48, and for *Agrobacterium*-mediated DNA transfer (Ishida et al. (1996) Nature Biotech  
30 14: 745-750).

Typically, plant transformation vectors include one or more cloned plant coding sequence (genomic or cDNA) under the transcriptional control of 5' and 3' regulatory sequences and a dominant selectable marker. Such plant transformation vectors typically also contain a promoter (e.g., a regulatory region controlling inducible or constitutive, environmentally-or

developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, an RNA processing signal (such as intron splice sites), a transcription termination site, and/or a polyadenylation signal.

5 Examples of constitutive plant promoters which can be useful for expressing the TF sequence include: the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant tissues (*see, e.g., Odel et al. (1985) Nature 313:810*); the nopaline synthase promoter (*An et al. (1988) Plant Physiol 88:547*); and the octopine synthase promoter (*Fromm et al. (1989) Plant Cell 1: 977*).

10 A variety of plant gene promoters that regulate gene expression in response to environmental, hormonal, chemical, developmental signals, and in a tissue-active manner can be used for expression of a TF sequence in plants. Choice of a promoter is based largely on the phenotype of interest and is determined by such factors as tissue (e.g., seed, fruit, root, pollen, vascular tissue, flower, carpel, etc.), inducibility (e.g., in response to wounding, heat, cold, drought, light, pathogens, etc.), timing, developmental stage, and the like. Numerous known  
15 promoters have been characterized and can favorably be employed to promote expression of a polynucleotide of the invention in a transgenic plant or cell of interest. For example, tissue specific promoters include: seed-specific promoters (such as the napin, phaseolin or DC3 promoter described in US Pat. No. 5,773,697), fruit-specific promoters that are active during fruit ripening (such as the *dru 1* promoter (US Pat. No. 5,783,393), or the 2A11 promoter (US Pat. No.  
20 4,943,674) and the tomato polygalacturonase promoter (*Bird et al. (1988) Plant Mol Biol 11:651*), root-specific promoters, such as those disclosed in US Patent Nos. 5,618,988, 5,837,848 and 5,905,186, pollen-active promoters such as PTA29, PTA26 and PTA13 (US Pat. No. 5,792,929), promoters active in vascular tissue (*Ringli and Keller (1998) Plant Mol Biol 37:977-988*), flower-specific (*Kaiser et al, (1995) Plant Mol Biol 28:231-243*), pollen (*Baerson et al. (1994) Plant Mol Biol 26:1947-1959*), carpels (*Ohl et al. (1990) Plant Cell 2:837-848*), pollen and ovules (*Baerson et al. (1993) Plant Mol Biol 22:255-267*), auxin-inducible promoters (such as that described in  
25 *van der Kop et al. (1999) Plant Mol Biol 39:979-990* or *Baumann et al. (1999) Plant Cell 11:323-334*), cytokinin-inducible promoter (*Guevara-Garcia (1998) Plant Mol Biol 38:743-753*), promoters responsive to gibberellin (*Shi et al. (1998) Plant Mol Biol 38:1053-1060*, *Willmott et al. (1998) 38:817-825*) and the like. Additional promoters are those that elicit expression in  
30 response to heat (*Ainley et al. (1993) Plant Mol Biol 22: 13-23*), light (e.g., the pea *rbcS-3A* promoter, *Kuhlemeier et al. (1989) Plant Cell 1:471*, and the maize *rbcS* promoter, *Schaffner and Sheen (1991) Plant Cell 3: 997*); wounding (e.g., *wun1*, *Siebertz et al. (1989) Plant Cell 1: 961*); pathogens (such as the PR-1 promoter described in *Buchel et al. (1999) Plant Mol. Biol. 40:387-*



396, and the PDF1.2 promoter described in Manners et al. (1998) Plant Mol. Biol. 38:1071-80), and chemicals such as methyl jasmonate or salicylic acid (Gatz et al. (1997) Plant Mol Biol 48: 89-108). In addition, the timing of the expression can be controlled by using promoters such as those acting at senescence (An and Amazon (1995) Science 270: 1986-1988); or late seed development (Odell et al. (1994) Plant Physiol 106:447-458).

Plant expression vectors can also include RNA processing signals that can be positioned within, upstream or downstream of the coding sequence. In addition, the expression vectors can include additional regulatory sequences from the 3'-untranslated region of plant genes, e.g., a 3' terminator region to increase mRNA stability of the mRNA, such as the PI-II terminator region of potato or the octopine or nopaline synthase 3' terminator regions.

#### Additional Expression Elements

Specific initiation signals can aid in efficient translation of coding sequences.

These signals can include, e.g., the ATG initiation codon and adjacent sequences. In cases where a coding sequence, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence (e.g., a mature protein coding sequence), or a portion thereof, is inserted, exogenous transcriptional control signals including the ATG initiation codon can be separately provided. The initiation codon is provided in the correct reading frame to facilitate transcription. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers appropriate to the cell system in use.

#### Expression Hosts

The present invention also relates to host cells which are transduced with vectors of the invention, and the production of polypeptides of the invention (including fragments thereof) by recombinant techniques. Host cells are genetically engineered (i.e., nucleic acids are introduced, e.g., transduced, transformed or transfected) with the vectors of this invention, which may be, for example, a cloning vector or an expression vector comprising the relevant nucleic acids herein. The vector is optionally a plasmid, a viral particle, a phage, a naked nucleic acids, *etc.* The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants, or amplifying the relevant gene. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to those skilled in the art and in the references cited herein, including, Sambrook and Ausubel.

The host cell can be a eukaryotic cell, such as a yeast cell, or a plant cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Plant protoplasts are also suitable for some applications. For example, the DNA fragments are introduced into plant tissues, cultured plant cells or plant protoplasts by standard methods including electroporation (Fromm et al., (1985) Proc. Natl. Acad. Sci. USA 82, 5824, infection by viral vectors such as cauliflower mosaic virus (CaMV) (Hohn et al., (1982) Molecular Biology of Plant Tumors, (Academic Press, New York) pp. 549-560; US 4,407,956), high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface (Klein et al., (1987) Nature 327, 70-73), use of pollen as vector (WO 85/01856), or use of *Agrobacterium tumefaciens* or *A. rhizogenes* carrying a T-DNA plasmid in which DNA fragments are cloned. The T-DNA plasmid is transmitted to plant cells upon infection by *Agrobacterium tumefaciens*, and a portion is stably integrated into the plant genome (Horsch et al. (1984) Science 233:496-498; Fraley et al. (1983) Proc. Natl. Acad. Sci. USA 80, 4803).

The cell can include a nucleic acid of the invention which encodes a polypeptide, wherein the cells expresses a polypeptide of the invention. The cell can also include vector sequences, or the like. Furthermore, cells and transgenic plants which include any polypeptide or nucleic acid above or throughout this specification, e.g., produced by transduction of a vector of the invention, are an additional feature of the invention.

For long-term, high-yield production of recombinant proteins, stable expression can be used. Host cells transformed with a nucleotide sequence encoding a polypeptide of the invention are optionally cultured under conditions suitable for the expression and recovery of the encoded protein from cell culture. The protein or fragment thereof produced by a recombinant cell may be secreted, membrane-bound, or contained intracellularly, depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides encoding mature proteins of the invention can be designed with signal sequences which direct secretion of the mature polypeptides through a prokaryotic or eukaryotic cell membrane.

#### Modified Amino Acids

Polypeptides of the invention may contain one or more modified amino acids.

The presence of modified amino acids may be advantageous in, for example, increasing polypeptide half-life, reducing polypeptide antigenicity or toxicity, increasing polypeptide storage stability, or the like. Amino acid(s) are modified, for example, co-translationally or post-translationally during recombinant production or modified by synthetic or chemical means.

Non-limiting examples of a modified amino acid include incorporation or other use of acetylated amino acids, glycosylated amino acids, sulfated amino acids, prenylated (e.g., farnesylated, geranylgeranylated) amino acids, PEG modified (e.g., "PEGylated") amino acids, biotinylated amino acids, carboxylated amino acids, phosphorylated amino acids, etc. References  
5 adequate to guide one of skill in the modification of amino acids are replete throughout the literature.

#### IDENTIFICATION OF ADDITIONAL FACTORS

A transcription factor provided by the present invention can also be used to identify additional endogenous or exogenous molecules that can affect a phenotype or trait of  
10 interest. On the one hand, such molecules include organic (small or large molecules) and/or inorganic compounds that affect expression of (i.e., regulate) a particular transcription factor. Alternatively, such molecules include endogenous molecules that are acted upon either at a transcriptional level by a transcription factor of the invention to modify a phenotype as desired. For example, the transcription factors can be employed to identify one or more downstream gene  
15 with which is subject to a regulatory effect of the transcription factor. In one approach, a transcription factor or transcription factor homologue of the invention is expressed in a host cell, e.g., a transgenic plant cell, tissue or explant, and expression products, either RNA or protein, of likely or random targets are monitored, e.g., by hybridization to a microarray of nucleic acid probes corresponding to genes expressed in a tissue or cell type of interest, by two-dimensional  
20 gel electrophoresis of protein products, or by any other method known in the art for assessing expression of gene products at the level of RNA or protein. Alternatively, a transcription factor of the invention can be used to identify promoter sequences (i.e., binding sites) involved in the regulation of a downstream target. After identifying a promoter sequence, interactions between the transcription factor and the promoter sequence can be modified by changing specific  
25 nucleotides in the promoter sequence or specific amino acids in the transcription factor that interact with the promoter sequence to alter a plant trait. Typically, transcription factor DNA binding sites are identified by gel shift assays. After identifying the promoter regions, the promoter region sequences can be employed in double-stranded DNA arrays to identify molecules that affect the interactions of the transcription factors with their promoters (Bulyk et al.  
30 (1999) Nature Biotechnology 17:573-577).

The identified transcription factors are also useful to identify proteins that modify the activity of the transcription factor. Such modification can occur by covalent modification, such as by phosphorylation, or by protein-protein (homo or-heteropolymer) interactions. Any

method suitable for detecting protein-protein interactions can be employed. Among the methods that can be employed are co-immunoprecipitation, cross-linking and co-purification through gradients or chromatographic columns, and the two-hybrid yeast system.

The two-hybrid system detects protein interactions *in vivo* and is described in Chien, et al., (1991), Proc. Natl. Acad. Sci. USA 88, 9578-9582 and is commercially available from Clontech (Palo Alto, Calif.). In such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to the TF polypeptide and the other consists of the transcription activator protein's activation domain fused to an unknown protein that is encoded by a cDNA that has been recombined into the plasmid as part of a cDNA library. The DNA-binding domain fusion plasmid and the cDNA library are transformed into a strain of the yeast *Saccharomyces cerevisiae* that contains a reporter gene (e.g., lacZ) whose regulatory region contains the transcription activator's binding site. Either hybrid protein alone cannot activate transcription of the reporter gene. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product. Then, the library plasmids responsible for reporter gene expression are isolated and sequenced to identify the proteins encoded by the library plasmids. After identifying proteins that interact with the transcription factors, assays for compounds that interfere with the TF protein-protein interactions can be preformed.

## 20 IDENTIFICATION OF MODULATORS

In addition to the intracellular molecules described above, extracellular molecules that alter activity or expression of a transcription factor, either directly or indirectly, can be identified. For example, the methods can entail first placing a candidate molecule in contact with a plant or plant cell. The molecule can be introduced by topical administration, such as spraying or soaking of a plant, and then the molecule's effect on the expression or activity of the TF polypeptide or the expression of the polynucleotide monitored. Changes in the expression of the TF polypeptide can be monitored by use of polyclonal or monoclonal antibodies, gel electrophoresis or the like. Changes in the expression of the corresponding polynucleotide sequence can be detected by use of microarrays, Northernblots, quantitative PCR, or any other technique for monitoring changes in mRNA expression. These techniques are exemplified in Ausubel et al. (eds) Current Protocols in Molecular Biology, John Wiley & Sons (1998). Such changes in the expression levels can be correlated with modified plant traits and thus identified

molecules can be useful for soaking or spraying on fruit, vegetable and grain crops to modify traits in plants.

Essentially any available composition can be tested for modulatory activity of expression or activity of any nucleic acid or polypeptide herein. Thus, available libraries of compounds such as chemicals, polypeptides, nucleic acids and the like can be tested for modulatory activity. Often, potential modulator compounds can be dissolved in aqueous or organic (e.g., DMSO-based) solutions for easy delivery to the cell or plant of interest in which the activity of the modulator is to be tested. Optionally, the assays are designed to screen large modulator composition libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays).

In one embodiment, high throughput screening methods involve providing a combinatorial library containing a large number of potential compounds (potential modulator compounds). Such "combinatorial chemical libraries" are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as target compounds.

A combinatorial chemical library can be, e.g., a collection of diverse chemical compounds generated by chemical synthesis or biological synthesis. For example, a combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (e.g., in one example, amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound of a set length). Exemplary libraries include peptide libraries, nucleic acid libraries, antibody libraries (see, e.g., Vaughn et al. (1996) Nature Biotechnology, 14(3):309-314 and PCT/US96/10287), carbohydrate libraries (see, e.g., Liang et al. Science (1996) 274:1520-1522 and U.S. Patent 5,593,853), peptide nucleic acid libraries (see, e.g., U.S. Patent 5,539,083), and small organic molecule libraries (see, e.g., benzodiazepines, Baum C&EN Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337) and the like.

Preparation and screening of combinatorial or other libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (see, e.g., U.S. Patent 5,010,175, Furka, Int. J. Pept. Prot. Res. 37:487-493 (1991) and Houghton et al. Nature 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used.

In addition, as noted, compound screening equipment for high-throughput screening is generally available, e.g., using any of a number of well known robotic systems that have also been developed for solution phase chemistries useful in assay systems. These systems include automated workstations including an automated synthesis apparatus and robotic systems utilizing robotic arms. Any of the above devices are suitable for use with the present invention, e.g., for high-throughput screening of potential modulators. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art.

Indeed, entire high throughput screening systems are commercially available.

These systems typically automate entire procedures including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. Similarly, microfluidic implementations of screening are also commercially available.

The manufacturers of such systems provide detailed protocols the various high throughput. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like. The integrated systems herein, in addition to providing for sequence alignment and, optionally, synthesis of relevant nucleic acids, can include such screening apparatus to identify modulators that have an effect on one or more polynucleotides or polypeptides according to the present invention.

In some assays it is desirable to have positive controls to ensure that the components of the assays are working properly. At least two types of positive controls are appropriate. That is, known transcriptional activators or inhibitors can be incubated with cells/plants/ etc. in one sample of the assay, and the resulting increase/decrease in transcription can be detected by measuring the resulting increase in RNA/ protein expression, etc., according to the methods herein. It will be appreciated that modulators can also be combined with transcriptional activators or inhibitors to find modulators which inhibit transcriptional activation or transcriptional repression. Either expression of the nucleic acids and proteins herein or any additional nucleic acids or proteins activated by the nucleic acids or proteins herein, or both, can be monitored.

In an embodiment, the invention provides a method for identifying compositions that modulate the activity or expression of a polynucleotide or polypeptide of the invention. For example, a test compound, whether a small or large molecule, is placed in contact with a cell,

plant (or plant tissue or explant), or composition comprising the polynucleotide or polypeptide of interest and a resulting effect on the cell, plant, (or tissue or explant) or composition is evaluated by monitoring, either directly or indirectly, one or more of: expression level of the polynucleotide or polypeptide, activity (or modulation of the activity) of the polynucleotide or polypeptide. In some cases, an alteration in a plant phenotype can be detected following contact of a plant (or plant cell, or tissue or explant) with the putative modulator, e.g., by modulation of expression or activity of a polynucleotide or polypeptide of the invention.

#### SUBSEQUENCES

Also contemplated are uses of polynucleotides, also referred to herein as oligonucleotides, typically having at least 12 bases, preferably at least 15, more preferably at least 20, 30, or 50 bases, which hybridize under at least highly stringent (or ultra-high stringent or ultra-ultra- high stringent conditions) conditions to a polynucleotide sequence described above. The polynucleotides may be used as probes, primers, sense and antisense agents, and the like, according to methods as noted *supra*.

Subsequences of the polynucleotides of the invention, including polynucleotide fragments and oligonucleotides are useful as nucleic acid probes and primers. An oligonucleotide suitable for use as a probe or primer is at least about 15 nucleotides in length, more often at least about 18 nucleotides, often at least about 21 nucleotides, frequently at least about 30 nucleotides, or about 40 nucleotides, or more in length. A nucleic acid probe is useful in hybridization protocols, e.g., to identify additional polypeptide homologues of the invention, including protocols for microarray experiments. Primers can be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR) or other nucleic-acid amplification methods. See Sambrook and Ausubel, *supra*.

In addition, the invention includes an isolated or recombinant polypeptide including a subsequence of at least about 15 contiguous amino acids encoded by the recombinant or isolated polynucleotides of the invention. For example, such polypeptides, or domains or fragments thereof, can be used as immunogens, e.g., to produce antibodies specific for the polypeptide sequence, or as probes for detecting a sequence of interest. A subsequence can range in size from about 15 amino acids in length up to and including the full length of the polypeptide.

## PRODUCTION OF TRANSGENIC PLANTS

### Modification of Traits

The polynucleotides of the invention are favorably employed to produce transgenic plants with various traits, or characteristics, that have been modified in a desirable manner, e.g., to improve the seed characteristics of a plant. For example, alteration of expression levels or patterns (e.g., spatial or temporal expression patterns) of one or more of the transcription factors (or transcription factor homologues) of the invention, as compared with the levels of the same protein found in a wild type plant, can be used to modify a plant's traits. An illustrative example of trait modification, modified structure and development characteristics, by altering expression levels of a particular transcription factor is described further in the Examples and the Sequence Listing.

### Antisense and Cosuppression Approaches

In addition to expression of the nucleic acids of the invention as gene replacement or plant phenotype modification nucleic acids, the nucleic acids are also useful for sense and anti-sense suppression of expression, e.g., to down-regulate expression of a nucleic acid of the invention, e.g., as a further mechanism for modulating plant phenotype. That is, the nucleic acids of the invention, or subsequences or anti-sense sequences thereof, can be used to block expression of naturally occurring homologous nucleic acids. A variety of sense and anti-sense technologies are known in the art, e.g., as set forth in Lichtenstein and Nellen (1997) Antisense Technology: A Practical Approach IRL Press at Oxford University, Oxford, England. In general, sense or anti-sense sequences are introduced into a cell, where they are optionally amplified, e.g., by transcription. Such sequences include both simple oligonucleotide sequences and catalytic sequences such as ribozymes.

For example, a reduction or elimination of expression (i.e., a "knock-out") of a transcription factor or transcription factor homologue polypeptide in a transgenic plant, e.g., to modify a plant trait, can be obtained by introducing an antisense construct corresponding to the polypeptide of interest as a cDNA. For antisense suppression, the transcription factor or homologue cDNA is arranged in reverse orientation (with respect to the coding sequence) relative to the promoter sequence in the expression vector. The introduced sequence need not be the full length cDNA or gene, and need not be identical to the cDNA or gene found in the plant type to be transformed. Typically, the antisense sequence need only be capable of hybridizing to the target gene or RNA of interest. Thus, where the introduced sequence is of shorter length, a higher degree of homology to the endogenous transcription factor sequence will be needed for effective antisense suppression. While antisense sequences of various lengths can be utilized, preferably,



the introduced antisense sequence in the vector will be at least 30 nucleotides in length, and improved antisense suppression will typically be observed as the length of the antisense sequence increases. Preferably, the length of the antisense sequence in the vector will be greater than 100 nucleotides. Transcription of an antisense construct as described results in the production of  
5 RNA molecules that are the reverse complement of mRNA molecules transcribed from the endogenous transcription factor gene in the plant cell.

Suppression of endogenous transcription factor gene expression can also be achieved using a ribozyme. Ribozymes are RNA molecules that possess highly specific endoribonuclease activity. The production and use of ribozymes are disclosed in U.S. Patent No.  
10 4,987,071 and U.S. Patent No. 5,543,508. Synthetic ribozyme sequences including antisense RNAs can be used to confer RNA cleaving activity on the antisense RNA, such that endogenous mRNA molecules that hybridize to the antisense RNA are cleaved, which in turn leads to an enhanced antisense inhibition of endogenous gene expression.

Vectors in which RNA encoded by a transcription factor or transcription factor  
15 homologue cDNA is over-expressed can also be used to obtain co-suppression of a corresponding endogenous gene, e.g., in the manner described in U.S. Patent No. 5,231,020 to Jorgensen. Such co-suppression (also termed sense suppression) does not require that the entire transcription factor cDNA be introduced into the plant cells, nor does it require that the introduced sequence be exactly identical to the endogenous transcription factor gene of interest. However, as with  
20 antisense suppression, the suppressive efficiency will be enhanced as specificity of hybridization is increased, e.g., as the introduced sequence is lengthened, and/or as the sequence similarity between the introduced sequence and the endogenous transcription factor gene is increased.

Vectors expressing an untranslatable form of the transcription factor mRNA, e.g., sequences comprising one or more stop codon, or nonsense mutation) can also be used to  
25 suppress expression of an endogenous transcription factor, thereby reducing or eliminating it's activity and modifying one or more traits. Methods for producing such constructs are described in U.S. Patent No. 5,583,021. Preferably, such constructs are made by introducing a premature stop codon into the transcription factor gene. Alternatively, a plant trait can be modified by gene silencing using double-strand RNA (Sharp (1999) Genes and Development 13: 139-141).

30 Another method for abolishing the expression of a gene is by insertion mutagenesis using the T-DNA of *Agrobacterium tumefaciens*. After generating the insertion mutants, the mutants can be screened to identify those containing the insertion in a transcription factor or transcription factor homologue gene. Plants containing a single transgene insertion

event at the desired gene can be crossed to generate homozygous plants for the mutation (Koncz et al. (1992) Methods in Arabidopsis Research, World Scientific).

Alternatively, a plant phenotype can be altered by eliminating an endogenous gene, such as a transcription factor or transcription factor homologue, e.g., by homologous recombination (Kempin et al. (1997) Nature 389:802).

A plant trait can also be modified by using the cre-lox system (for example, as described in US Pat. No. 5,658,772). A plant genome can be modified to include first and second lox sites that are then contacted with a Cre recombinase. If the lox sites are in the same orientation, the intervening DNA sequence between the two sites is excised. If the lox sites are in the opposite orientation, the intervening sequence is inverted.

The polynucleotides and polypeptides of this invention can also be expressed in a plant in the absence of an expression cassette by manipulating the activity or expression level of the endogenous gene by other means. For example, by ectopically expressing a gene by T-DNA activation tagging (Ichikawa et al. (1997) Nature 390 698-701; Kakimoto et al. (1996) Science 274: 982-985). This method entails transforming a plant with a gene tag containing multiple transcriptional enhancers and once the tag has inserted into the genome, expression of a flanking gene coding sequence becomes deregulated. In another example, the transcriptional machinery in a plant can be modified so as to increase transcription levels of a polynucleotide of the invention (See, e.g., PCT Publications WO 96/06166 and WO 98/53057 which describe the modification of the DNA binding specificity of zinc finger proteins by changing particular amino acids in the DNA binding motif).

The transgenic plant can also include the machinery necessary for expressing or altering the activity of a polypeptide encoded by an endogenous gene, for example by altering the phosphorylation state of the polypeptide to maintain it in an activated state.

Transgenic plants (or plant cells, or plant explants, or plant tissues) incorporating the polynucleotides of the invention and/or expressing the polypeptides of the invention can be produced by a variety of well established techniques as described above. Following construction of a vector, most typically an expression cassette, including a polynucleotide, e.g., encoding a transcription factor or transcription factor homologue, of the invention, standard techniques can be used to introduce the polynucleotide into a plant, a plant cell, a plant explant or a plant tissue of interest. Optionally, the plant cell, explant or tissue can be regenerated to produce a transgenic plant.

The plant can be any higher plant, including gymnosperms, monocotyledonous and dicotyledonous plants. Suitable protocols are available for *Leguminosae* (alfalfa, soybean,

clover, etc.), *Umbelliferae* (carrot, celery, parsnip), *Cruciferae* (cabbage, radish, rapeseed, broccoli, etc.), *Curcubitaceae* (melons and cucumber), *Gramineae* (wheat, corn, rice, barley, millet, etc.), *Solanaceae* (potato, tomato, tobacco, peppers, etc.), and various other crops. See protocols described in Ammirato et al. (1984) Handbook of Plant Cell Culture –Crop Species.

- 5 Macmillan Publ. Co. Shimamoto et al. (1989) Nature 338:274-276; Fromm et al. (1990) Bio/Technology 8:833-839; and Vasil et al. (1990) Bio/Technology 8:429-434.

Transformation and regeneration of both monocotyledonous and dicotyledonous plant cells is now routine, and the selection of the most appropriate transformation technique will be determined by the practitioner. The choice of method will vary with the type of plant to be  
10 transformed; those skilled in the art will recognize the suitability of particular methods for given plant types. Suitable methods can include, but are not limited to: electroporation of plant protoplasts; liposome-mediated transformation; polyethylene glycol (PEG) mediated transformation; transformation using viruses; micro-injection of plant cells; micro-projectile bombardment of plant cells; vacuum infiltration; and *Agrobacterium tumefaciens* mediated  
15 transformation. Transformation means introducing a nucleotide sequence in a plant in a manner to cause stable or transient expression of the sequence.

Successful examples of the modification of plant characteristics by transformation with cloned sequences which serve to illustrate the current knowledge in this field of technology, and which are herein incorporated by reference, include: U.S. Patent Nos.  
20 5,571,706; 5,677,175; 5,510,471; 5,750,386; 5,597,945; 5,589,615; 5,750,871; 5,268,526; 5,780,708; 5,538,880; 5,773,269; 5,736,369 and 5,610,042.

Following transformation, plants are preferably selected using a dominant selectable marker incorporated into the transformation vector. Typically, such a marker will confer antibiotic or herbicide resistance on the transformed plants, and selection of transformants  
25 can be accomplished by exposing the plants to appropriate concentrations of the antibiotic or herbicide.

After transformed plants are selected and grown to maturity, those plants showing a modified trait are identified. The modified trait can be any of those traits described above. Additionally, to confirm that the modified trait is due to changes in expression levels or  
30 activity of the polypeptide or polynucleotide of the invention can be determined by analyzing mRNA expression using Northern blots, RT-PCR or microarrays, or protein expression using immunoblots or Western blots or gel shift assays.

### INTEGRATED SYSTEMS—SEQUENCE IDENTITY

Additionally, the present invention may be an integrated system, computer or computer readable medium that comprises an instruction set for determining the identity of one or more sequences in a database. In addition, the instruction set can be used to generate or identify sequences that meet any specified criteria. Furthermore, the instruction set may be used to associate or link certain functional benefits, such modified structure and development characteristics, with one or more identified sequence.

For example, the instruction set can include, e.g., a sequence comparison or other alignment program, e.g., an available program such as, for example, the Wisconsin Package Version 10.0, such as BLAST, FASTA, PILEUP, FINDPATTERNS or the like (GCG, Madison, WI). Public sequence databases such as GenBank, EMBL, Swiss-Prot and PIR or private sequence databases such as PhytoSeq (Incyte Pharmaceuticals, Palo Alto, CA) can be searched.

Alignment of sequences for comparison can be conducted by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman (1988) Proc. Natl. Acad. Sci. U.S.A. 85: 2444, by computerized implementations of these algorithms. After alignment, sequence comparisons between two (or more) polynucleotides or polypeptides are typically performed by comparing sequences of the two sequences over a comparison window to identify and compare local regions of sequence similarity. The comparison window can be a segment of at least about 20 contiguous positions, usually about 50 to about 200, more usually about 100 to about 150 contiguous positions. A description of the method is provided in Ausubel et al., *supra*.

A variety of methods of determining sequence relationships can be used, including manual alignment and computer assisted sequence alignment and analysis. This later approach is a preferred approach in the present invention, due to the increased throughput afforded by computer assisted methods. As noted above, a variety of computer programs for performing sequence alignment are available, or can be produced by one of skill.

One example algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul et al. J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available, e.g., through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is

referred to as the neighborhood word score threshold (Altschul et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for  
5 nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of  
10 one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E)  
15 of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see*, e.g., Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5787). One measure of similarity provided  
20 by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence (and, therefore, in this context, homologous) if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, or less than about 0.01, and or  
25 even less than about 0.001. An additional example of a useful sequence alignment algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. The program can align, e.g., up to 300 sequences of a maximum length of 5,000 letters.

The integrated system, or computer typically includes a user input interface  
30 allowing a user to selectively view one or more sequence records corresponding to the one or more character strings, as well as an instruction set which aligns the one or more character strings with each other or with an additional character string to identify one or more region of sequence similarity. The system may include a link of one or more character strings with a particular

phenotype or gene function. Typically, the system includes a user readable output element which displays an alignment produced by the alignment instruction set.

The methods of this invention can be implemented in a localized or distributed computing environment. In a distributed environment, the methods may implemented on a single computer comprising multiple processors or on a multiplicity of computers. The computers can be linked, e.g. through a common bus, but more preferably the computer(s) are nodes on a network. The network can be a generalized or a dedicated local or wide-area network and, in certain preferred embodiments, the computers may be components of an intra-net or an internet.

Thus, the invention provides methods for identifying a sequence similar or homologous to one or more polynucleotides as noted herein, or one or more target polypeptides encoded by the polynucleotides, or otherwise noted herein and may include linking or associating a given plant phenotype or gene function with a sequence. In the methods, a sequence database is provided (locally or across an inter or intra net) and a query is made against the sequence database using the relevant sequences herein and associated plant phenotypes or gene functions.

Any sequence herein can be entered into the database, before or after querying the database. This provides for both expansion of the database and, if done before the querying step, for insertion of control sequences into the database. The control sequences can be detected by the query to ensure the general integrity of both the database and the query. As noted, the query can be performed using a web browser based interface. For example, the database can be a centralized public database such as those noted herein, and the querying can be done from a remote terminal or computer across an internet or intranet.

### EXAMPLES

The following examples are intended to illustrate but not limit the present invention.

#### EXAMPLE I. FULL LENGTH GENE IDENTIFICATION AND CLONING

Putative transcription factor sequences (genomic or ESTs) related to known transcription factors were identified in the *Arabidopsis thaliana* GenBank database using the tblastn sequence analysis program using default parameters and a P-value cutoff threshold of  $-4$  or  $-5$  or lower, depending on the length of the query sequence. Putative transcription factor sequence hits were then screened to identify those containing particular sequence strings. If the sequence hits contained such sequence strings, the sequences were confirmed as transcription factors.

Alternatively, *Arabidopsis thaliana* cDNA libraries derived from different tissues or treatments, or genomic libraries were screened to identify novel members of a transcription family using a low stringency hybridization approach. Probes were synthesized using gene specific primers in a standard PCR reaction (annealing temperature 60° C) and labeled with <sup>32</sup>P dCTP using the High Prime DNA Labeling Kit (Boehringer Mannheim). Purified radiolabelled probes were added to filters immersed in Church hybridization medium (0.5 M NaPO<sub>4</sub> pH 7.0, 7% SDS, 1 % w/v bovine serum albumin) and hybridized overnight at 60 °C with shaking. Filters were washed two times for 45 to 60 minutes with 1xSSC, 1% SDS at 60° C.

To identify additional sequence 5' or 3' of a partial cDNA sequence in a cDNA library, 5' and 3' rapid amplification of cDNA ends (RACE) was performed using the Marathon™ cDNA amplification kit (Clontech, Palo Alto, CA). Generally, the method entailed first isolating poly(A) mRNA, performing first and second strand cDNA synthesis to generate double stranded cDNA, blunting cDNA ends, followed by ligation of the Marathon™ Adaptor to the cDNA to form a library of adaptor-ligated ds cDNA.

Gene-specific primers were designed to be used along with adaptor specific primers for both 5' and 3' RACE reactions. Nested primers, rather than single primers, were used to increase PCR specificity. Using 5' and 3' RACE reactions, 5' and 3' RACE fragments were obtained, sequenced and cloned. The process can be repeated until 5' and 3' ends of the full-length gene were identified. Then the full-length cDNA was generated by PCR using primers specific to 5' and 3' ends of the gene by end-to-end PCR.

#### EXAMPLE II. CONSTRUCTION OF EXPRESSION VECTORS

The sequence was amplified from a genomic or cDNA library using primers specific to sequences upstream and downstream of the coding region. The expression vector was pMEN20 or pMEN65, which are both derived from pMON316 (Sanders et al, (1987) Nucleic Acids Research 15:1543-58) and contain the CaMV 35S promoter to express transgenes. To clone the sequence into the vector, both pMEN20 and the amplified DNA fragment were digested separately with SalI and NotI restriction enzymes at 37° C for 2 hours. The digestion products were subject to electrophoresis in a 0.8% agarose gel and visualized by ethidium bromide staining. The DNA fragments containing the sequence and the linearized plasmid were excised and purified by using a Qiaquick gel extraction kit (Qiagen, CA). The fragments of interest were ligated at a ratio of 3:1 (vector to insert). Ligation reactions using T4 DNA ligase (New England Biolabs, MA) were carried out at 16° C for 16 hours. The ligated DNAs were transformed into

competent cells of the *E. coli* strain DH5alpha by using the heat shock method. The transformations were plated on LB plates containing 50 mg/l kanamycin (Sigma).

Individual colonies were grown overnight in five milliliters of LB broth containing 50 mg/l kanamycin at 37° C. Plasmid DNA was purified by using Qiaquick Mini  
5 Prep kits (Qiagen, CA).

#### EXAMPLE III. TRANSFORMATION OF *AGROBACTERIUM* WITH THE EXPRESSION VECTOR

After the plasmid vector containing the gene was constructed, the vector was used to transform *Agrobacterium tumefaciens* cells expressing the gene products. The stock of  
10 *Agrobacterium tumefaciens* cells for transformation were made as described by Nagel et al. (1990) FEMS Microbiol Letts. 67: 325-328. *Agrobacterium* strain ABI was grown in 250 ml LB medium (Sigma) overnight at 28°C with shaking until an absorbance ( $A_{600}$ ) of 0.5 – 1.0 was reached. Cells were harvested by centrifugation at 4,000 x g for 15 min at 4° C. Cells were then resuspended in 250 µl chilled buffer (1 mM HEPES, pH adjusted to 7.0 with KOH). Cells were  
15 centrifuged again as described above and resuspended in 125 µl chilled buffer. Cells were then centrifuged and resuspended two more times in the same HEPES buffer as described above at a volume of 100 µl and 750 µl, respectively. Resuspended cells were then distributed into 40 µl aliquots, quickly frozen in liquid nitrogen, and stored at -80° C.

*Agrobacterium* cells were transformed with plasmids prepared as described  
20 above following the protocol described by Nagel et al. For each DNA construct to be transformed, 50 – 100 ng DNA (generally resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0) was mixed with 40 µl of *Agrobacterium* cells. The DNA/cell mixture was then transferred to a chilled cuvette with a 2mm electrode gap and subject to a 2.5 kV charge dissipated at 25 µF and 200 µF using a Gene Pulser II apparatus (Bio-Rad). After electroporation, cells were  
25 immediately resuspended in 1.0 ml LB and allowed to recover without antibiotic selection for 2 – 4 hours at 28° C in a shaking incubator. After recovery, cells were plated onto selective medium of LB broth containing 100 µg/ml spectinomycin (Sigma) and incubated for 24-48 hours at 28° C. Single colonies were then picked and inoculated in fresh medium. The presence of the plasmid construct was verified by PCR amplification and sequence analysis.

#### EXAMPLE IV. TRANSFORMATION OF *ARABIDOPSIS* PLANTS WITH *AGROBACTERIUM TUMEFACIENS* WITH EXPRESSION VECTOR

After transformation of *Agrobacterium tumefaciens* with plasmid vectors containing the gene, single *Agrobacterium* colonies were identified, propagated, and used to



transform *Arabidopsis* plants. Briefly, 500 ml cultures of LB medium containing 50 mg/l kanamycin were inoculated with the colonies and grown at 28° C with shaking for 2 days until an absorbance ( $A_{600}$ ) of > 2.0 is reached. Cells were then harvested by centrifugation at 4,000 x g for 10 min, and resuspended in infiltration medium (1/2 X Murashige and Skoog salts (Sigma), 1 X Gamborg's B-5 vitamins (Sigma), 5.0% (w/v) sucrose (Sigma), 0.044  $\mu$ M benzylamino purine (Sigma), 200  $\mu$ l/L Silwet L-77 (Lehle Seeds) until an absorbance ( $A_{600}$ ) of 0.8 was reached.

Prior to transformation, *Arabidopsis thaliana* seeds (ecotype Columbia) were sown at a density of ~10 plants per 4" pot onto Pro-Mix BX potting medium (Hummert International) covered with fiberglass mesh (18 mm X 16 mm). Plants were grown under continuous illumination (50-75  $\mu$ E/m<sup>2</sup>/sec) at 22-23° C with 65-70% relative humidity. After about 4 weeks, primary inflorescence stems (bolts) are cut off to encourage growth of multiple secondary bolts. After flowering of the mature secondary bolts, plants were prepared for transformation by removal of all siliques and opened flowers.

The pots were then immersed upside down in the mixture of *Agrobacterium* infiltration medium as described above for 30 sec, and placed on their sides to allow draining into a 1' x 2' flat surface covered with plastic wrap. After 24 h, the plastic wrap was removed and pots are turned upright. The immersion procedure was repeated one week later, for a total of two immersions per pot. Seeds were then collected from each transformation pot and analyzed following the protocol described below.

#### EXAMPLE V. IDENTIFICATION OF ARABIDOPSIS PRIMARY TRANSFORMANTS

Seeds collected from the transformation pots were sterilized essentially as follows. Seeds were dispersed into in a solution containing 0.1% (v/v) Triton X-100 (Sigma) and sterile H<sub>2</sub>O and washed by shaking the suspension for 20 min. The wash solution was then drained and replaced with fresh wash solution to wash the seeds for 20 min with shaking. After removal of the second wash solution, a solution containing 0.1% (v/v) Triton X-100 and 70% ethanol (Equistar) was added to the seeds and the suspension was shaken for 5 min. After removal of the ethanol/detergent solution, a solution containing 0.1% (v/v) Triton X-100 and 30% (v/v) bleach (Clorox) was added to the seeds, and the suspension was shaken for 10 min. After removal of the bleach/detergent solution, seeds were then washed five times in sterile distilled H<sub>2</sub>O. The seeds were stored in the last wash water at 4° C for 2 days in the dark before being plated onto antibiotic selection medium (1 X Murashige and Skoog salts (pH adjusted to 5.7 with 1M KOH), 1 X Gamborg's B-5 vitamins, 0.9% phytagar (Life Technologies), and 50 mg/l kanamycin). Seeds were germinated under continuous illumination (50-75  $\mu$ E/m<sup>2</sup>/sec) at 22-23°

C. After 7-10 days of growth under these conditions, kanamycin resistant primary transformants (T<sub>1</sub> generation) were visible and obtained. These seedlings were transferred first to fresh selection plates where the seedlings continued to grow for 3-5 more days, and then to soil (Pro-Mix BX potting medium).

Primary transformants were crossed and progeny seeds (T<sub>2</sub>) collected; kanamycin resistant seedlings were selected and analyzed. The expression levels of the recombinant polynucleotides in the transformants varies from about a 5% expression level increase to a least a 100% expression level increase. Similar observations are made with respect to polypeptide level expression.

#### EXAMPLE VI. IDENTIFICATION OF ARABIDOPSIS PLANTS WITH TRANSCRIPTION FACTOR GENE KNOCKOUTS

The screening of insertion mutagenized *Arabidopsis* collections for null mutants in a known target gene was essentially as described in Krysan et al (1999) Plant Cell 11:2283-2290. Briefly, gene-specific primers, nested by 5-250 pb to each others, were designed from the 5' and 3' regions of a known target gene. Similarly, nested sets of primers were also created specific to each of the T-DNA or transposon ends (the "right" and "left" borders). All possible combinations of gene specific and T-DNA/transposon primers were used to detect by PCR an insertion event within or close to the target gene. The amplified DNA fragments were then sequenced which allows the precise determination of the T-DNA/transposon insertion point relative to the target gene. Insertion events within the coding or intervening sequence of the genes were deconvoluted from a pool comprising a plurality of insertion events to a single unique mutant plant for functional characterization. The method is described in more detail in Yu and Adam, US Application Serial No. 09/177,733 filed October 23, 1998.

#### EXAMPLE VII. IDENTIFICATION OF STRUCTURE AND DEVELOPMENT CHARACTERISTICS PHENOTYPE IN OVEREXPRESSOR OR GENE KNOCKOUT PLANTS

Experiments were performed to identify those transformants or knockouts that exhibited a modified structure and development characteristics. For such studies, the transformants were observed by eye to identify novel structural or developmental characteristics associated with the ectopic expression of the polynucleotides or polypeptides of the invention.

Table 3 shows the phenotypes observed for particular overexpressor or knockout plants and provides the SEQ ID No., the internal reference code (GID), whether a knockout or overexpressor plant was analyzed and the observed phenotype.

**Table 3**

SEQ ID No.	GID	Knockout (KO) or overexpressor (KO)	Phenotype observed
1	G727	OE	Plants were small, and more dark green in color, late flowering and poorly fertile.
3	G732	OE	Plants were small and inflorescence was unelongated. Flowers parts appeared to be unelongated and the plants were semi-sterile.
5	G9	OE	Increased root mass
7	G428	OE	Lobed and highly serrated leaves and abnormal first and second whorl floral organs
9	G869	OE	Undeveloped or small anthers
11	G1269	OE	Extended petioles and leaves pointed upwards
13	G1038	OE	Altered leaf shape
15	G438	KO	Reduced lignin in stem
17	G571	KO	Delayed senescence at the end of the plant lifecycle
19	G748	OE	More vascular bundles in stem
21	G431	OE	Severe developmental abnormalities such as altered branching, twisted rosette leaves, flowers with missing pistils, fused stamens and atypical numbers of petals and stamens, reduced secondary bolts, and lack of cauline leaves.
23	G187	OE	Plants had long, thin cotyledons and reduced apical dominance. Several flower abnormalities, including underdeveloped, sepaloid petals and underdeveloped anthers were also observed.
25	G470	OE	Plants were sterile due to failure of anthers to elongate
27	G615	OE	Plants were sterile due to failure of anthers to develop and failure of stamens to elongate. Fused cotyledons and absence of a shoot apical meristem and true leaves was also observed.
29	G1073	OE	Increased plant size and serrated leaves

For a particular overexpressor that shows a less beneficial structure and development characteristic, it may be more useful to select a plant with a decreased expression of the particular transcription factor. For a particular knockout that shows a less beneficial structure and development characteristic, it may be more useful to select a plant with an increased expression of the particular transcription factor.

EXAMPLE VIII. IDENTIFICATION OF HOMOLOGOUS SEQUENCES

Homologous sequences from *Arabidopsis* and plant species other than *Arabidopsis* were identified using database sequence search tools, such as the Basic Local Alignment Search Tool (BLAST) (Altschul et al. (1990) J. Mol. Biol. 215:403-410; and Altschul et al. (1997) Nucl. Acid Res. 25: 3389-3402). The tblastx sequence analysis programs were employed using the BLOSUM-62 scoring matrix (Henikoff, S. and Henikoff, J. G. (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919).

Identified *Arabidopsis* homologous sequences are provided in Figure 2 and included in the Sequence Listing. The percent sequence identity among these sequences is as low as 47% sequence identity. Additionally, the entire NCBI GenBank database was filtered for sequences from all plants except *Arabidopsis thaliana* by selecting all entries in the NCBI GenBank database associated with NCBI taxonomic ID 33090 (Viridiplantae; all plants) and excluding entries associated with taxonomic ID 3701 (*Arabidopsis thaliana*). These sequences were compared to sequences representing genes of SEQ IDs Nos. 1-54 on 9/26/2000 using the Washington University TBLASTX algorithm (version 2.0a19MP). For each gene of SEQ IDs Nos. 1-54, individual comparisons were ordered by probability score (P-value), where the score reflects the probability that a particular alignment occurred by chance. For example, a score of  $3.6e-40$  is  $3.6 \times 10^{-40}$ . For up to ten species, the gene with the lowest P-value (and therefore the most likely homolog) is listed in Figure 3.

In addition to P-values, comparisons were also scored by percentage identity. Percentage identity reflects the degree to which two segments of DNA or protein are identical over a particular length. The ranges of percent identity between the non-*Arabidopsis* genes shown in Figure 3 and the *Arabidopsis* genes in the sequence listing are: SEQ ID No. 1: 36%-69%; SEQ ID No. 3: 46%-54%; SEQ ID No. 5: 57%-72%; SEQ ID No. 7: 54%-69%; SEQ ID No. 9: 31%-68%; SEQ ID No. 11: 47%-90%; SEQ ID No. 13: 34%-82%; SEQ ID No. 15: 49%-88%; SEQ ID No. 17: 56%-67%; SEQ ID No. 19: 39%-61%; SEQ ID No. 21: 61%-87%; SEQ ID No. 23: 38%-85%; SEQ ID No. 25: 44%-94%; SEQ ID No. 27: 35%-44%; SEQ ID No. 29: 37%-71%; SEQ ID No. 31: 38%-77%; SEQ ID No. 33: 57%-69%; SEQ ID No. 35: 54%-69%; SEQ ID No. 37: 60%-75%; SEQ ID No. 39: 47%-65%; SEQ ID No. 41: 60%-88%; SEQ ID No. 43: 43%-87%; and SEQ ID No. 45: 53%-97%.

The polynucleotides and polypeptides in the Sequence Listing and the identified homologous sequences may be stored in a computer system and have associated or linked with the sequences a function, such as that the polynucleotides and polypeptides are useful for modifying the structure and development characteristics of a plant.

All references, publications, patents and other documents herein are incorporated by reference in their entirety for all purposes. Although the invention has been described with reference to the embodiments and examples above, it should be understood that various  
5 modifications can be made without departing from the spirit of the invention.

What is claimed is:

1. A transgenic plant with modified structure and development characteristics, which plant comprises a recombinant polynucleotide comprising a nucleotide sequence selected from the group consisting of:
  - 5 (a) a nucleotide sequence encoding a polypeptide comprising a sequence selected from SEQ ID Nos. 2N, where N=1-23, or a complementary nucleotide sequence thereof;
  - (b) a nucleotide sequence encoding a polypeptide comprising a conservatively substituted variant of a polypeptide of (a);
  - (c) a nucleotide sequence comprising a sequence selected from those of SEQ ID Nos. 2N-1, where N=1-23, or a complementary nucleotide sequence thereof;
  - 10 (d) a nucleotide sequence comprising silent substitutions in a nucleotide sequence of (c);
  - (e) a nucleotide sequence which hybridizes under stringent conditions to a nucleotide sequence of one or more of: (a), (b), (c), or (d);
  - (f) a nucleotide sequence comprising at least 15 consecutive nucleotides of a sequence of any of (a)-(e);
  - 15 (g) a nucleotide sequence comprising a subsequence or fragment of any of (a)-(f), which subsequence or fragment encodes a polypeptide that modifies a plant's structure and development characteristics;
  - (h) a nucleotide sequence having at least 31% sequence identity to a nucleotide sequence of any of (a)-(g);
  - 20 (i) a nucleotide sequence having at least 60% identity sequence identity to a nucleotide sequence of any of (a)-(g);
  - (j) a nucleotide sequence which encodes a polypeptide having at least 31% identity sequence identity to a polypeptide of SEQ ID Nos. 2N, where N=1-23;
  - 25 (k) a nucleotide sequence which encodes a polypeptide having at least 60% identity sequence identity to a polypeptide of SEQ ID Nos. 2N, where N=1-23; and
  - (l) a nucleotide sequence which encodes a polypeptide having at least 65% sequence identity to a conserved domain of a polypeptide of SEQ ID Nos. 2N, where N=1-23.
- 30 2. The transgenic plant of claim 1, further comprising a constitutive, inducible, or tissue-active promoter operably linked to said nucleotide sequence.
3. The transgenic plant of claim 1, wherein the plant is selected from the group consisting of: soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugarcane, turf,

banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honeydew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits, and vegetable brassicas.

5

4. An isolated or recombinant polynucleotide comprising a nucleotide sequence selected from the group consisting of:

- (a) a nucleotide sequence encoding a polypeptide comprising a sequence selected from SEQ ID Nos. 2N, where N=1-23, or a complementary nucleotide sequence thereof;
- 10 (b) a nucleotide sequence encoding a polypeptide comprising a conservatively substituted variant of a polypeptide of (a);
- (c) a nucleotide sequence comprising a sequence selected from those of SEQ ID Nos. 2N-1, where N=1-23, or a complementary nucleotide sequence thereof;
- (d) a nucleotide sequence comprising silent substitutions in a nucleotide sequence of (c);
- 15 (e) a nucleotide sequence which hybridizes under stringent conditions to a nucleotide sequence of one or more of: (a), (b), (c), or (d);
- (f) a nucleotide sequence comprising at least 15 consecutive nucleotides of a sequence of any of (a)-(e);
- (g) a nucleotide sequence comprising a subsequence or fragment of any of (a)-(f), which
- 20 subsequence or fragment encodes a polypeptide that modifies a plant's structure and development characteristics;
- (h) a nucleotide sequence having at least 31% sequence identity to a nucleotide sequence of any of (a)-(g);
- (i) a nucleotide sequence having at least 60% identity sequence identity to a nucleotide
- 25 sequence of any of (a)-(g);
- (j) a nucleotide sequence which encodes a polypeptide having at least 31% identity sequence identity to a polypeptide of SEQ ID Nos. 2N, where N=1-23;
- (k) a nucleotide sequence which encodes a polypeptide having at least 60% identity sequence identity to a polypeptide of SEQ ID Nos. 2N, where N=1-23; and
- 30 (l) a nucleotide sequence which encodes a conserved domain of a polypeptide having at least 65% sequence identity to a conserved domain of a polypeptide of SEQ ID Nos. 2N, where N=1-23.

5. The isolated or recombinant polynucleotide of claim 4, further comprising a constitutive, inducible, or tissue-active promoter operably linked to the nucleotide sequence.

6. A cloning or expression vector comprising the isolated or recombinant polynucleotide of claim 4.

7. A cell comprising the cloning or expression vector of claim 6.

8. A transgenic plant comprising the isolated or recombinant polynucleotide of claim 4.

9. A composition produced by one or more of:

- (a) incubating one or more polynucleotide of claim 4 with a nuclease;
- (b) incubating one or more polynucleotide of claim 4 with a restriction enzyme;
- (c) incubating one or more polynucleotide of claim 4 with a polymerase;
- (d) incubating one or more polynucleotide of claim 4 with a polymerase and a primer;
- (e) incubating one or more polynucleotide of claim 4 with a cloning vector, or
- (f) incubating one or more polynucleotide of claim 4 with a cell.

10. A composition comprising two or more different polynucleotides of claim 4.

11. An isolated or recombinant polypeptide comprising a subsequence of at least about 15 contiguous amino acids encoded by the recombinant or isolated polynucleotide of claim 4.

12. A plant ectopically expressing an isolated polypeptide of claim 11.

13. A method for producing a plant having a modified structure and development characteristic, the method comprising altering the expression of the isolated or recombinant polynucleotide of claim 4 or the expression levels or activity of a polypeptide of claim 11 in a plant, thereby producing a modified plant, and selecting the modified plant for modified structure and development characteristics thereby providing the modified plant with a modified structure and development characteristics.

14. The method of claim 13, wherein the polynucleotide is a polynucleotide of claim 4.



15. A method of identifying a factor that is modulated by or interacts with a polypeptide encoded by a polynucleotide of claim 4, the method comprising:

- (a) expressing a polypeptide encoded by the polynucleotide in a plant; and
- (b) identifying at least one factor that is modulated by or interacts with the polypeptide.

5

16. The method of claim 15, wherein the identifying is performed by detecting binding by the polypeptide to a promoter sequence, or detecting interactions between an additional protein and the polypeptide in a yeast two hybrid system.

10 17. The method of claim 15, wherein the identifying is performed by detecting expression of a factor by hybridization to a microarray, subtractive hybridization or differential display.

18. A method of identifying a molecule that modulates activity or expression of a polynucleotide or polypeptide of interest, the method comprising:

- 15 (a) placing the molecule in contact with a plant comprising the polynucleotide or polypeptide encoded by the polynucleotide of claim 4; and,
- (b) monitoring one or more of:
  - (i) expression level of the polynucleotide in the plant;
  - (ii) expression level of the polypeptide in the plant;
  - 20 (iii) modulation of an activity of the polypeptide in the plant; or
  - (iv) modulation of an activity of the polynucleotide in the plant.

19. An integrated system, computer or computer readable medium comprising one or more character strings corresponding to a polynucleotide of claim 4, or to a polypeptide encoded by the polynucleotide.

25

20. The integrated system, computer or computer readable medium of claim 19, further comprising a link between said one or more sequence strings to a modified plant structure and development characteristics phenotype.

30

21. A method of identifying a sequence similar or homologous to one or more polynucleotides of claim 4, or one or more polypeptides encoded by the polynucleotides, the method comprising:

- (a) providing a sequence database; and,

(b) querying the sequence database with one or more target sequences corresponding to the one or more polynucleotides or to the one or more polypeptides to identify one or more sequence members of the database that display sequence similarity or homology to one or more of the one or more target sequences.

5

22. The method of claim 21, wherein the querying comprises aligning one or more of the target sequences with one or more of the one or more sequence members in the sequence database.

10

23. The method of claim 21, wherein the querying comprises identifying one or more of the one or more sequence members of the database that meet a user-selected identity criteria with one or more of the target sequences.

15

24. The method of claim 21, further comprising linking the one or more of the polynucleotides of claim 4, or encoded polypeptides, to a modified plant structure and development characteristics phenotype.

20

25. A plant comprising altered expression levels of an isolated or recombinant polynucleotide of claim 4.

26. A plant comprising altered expression levels or the activity of an isolated or recombinant polypeptide of claim 11.

25

27. A plant lacking a nucleotide sequence encoding a polypeptide of claim 11.

Figure 1

SEQ ID No.	GID	cDNA or protein	conserved domain
1	G727	cDNA	
2	G727	protein	226-269
3	G732	cDNA	
4	G732	protein	31-9
5	G9	cDNA	
6	G9	protein	62-127
7	G428	cDNA	
8	G428	protein	229-292
9	G869	cDNA	
10	G869	protein	109-177
11	G1269	cDNA	
12	G1269	protein	27-83
13	G1038	cDNA	
14	G1038	protein	198-247
15	G438	cDNA	
16	G438	protein	22-85
17	G571	cDNA	
18	G571	protein	160-220
19	G748	cDNA	
20	G748	protein	112-140
21	G431	cDNA	
22	G431	protein	286-335
23	G187	cDNA	
24	G187	protein	172-228
25	G470	cDNA	
26	G470	protein	61-393
27	G615	cDNA	
28	G615	protein	88-147
29	G1073	cDNA	
30	G1073	protein	33-42, 78-175

Figure 2

SEQ ID No.	GID	homolog	cDNA or protein	conserved domain
31	G1493	homolog of G727	cDNA	
32	G1493	homolog of G727	protein	242-289
33	G993	homolog of G9	cDNA	
34	G993	homolog of G9	protein	69-134
35	G867	homolog of G9	cDNA	
36	G867	homolog of G9	protein	59-124
37	G1930	homolog of G9	cDNA	
38	G1930	homolog of G9	protein	59-124
39	G1594	homolog of G428	cDNA	
40	G1594	homolog of G428	protein	262-325
41	G391	homolog of G438	cDNA	
42	G391	homolog of G438	protein	25-85
43	G390	homolog of G438	cDNA	
44	G390	homolog of G438	protein	18-81
45	G1548	homolog of G438	cDNA	
46	G1548	homolog of G438	protein	17-77

Figure 3A

SEQ ID No.	GID	Genbank NID	P-value	Species
1	G727	7283684	2.20E-56	Glycine max
1	G727	7206180	8.40E-42	Medicago truncatula
1	G727	7614196	2.20E-40	Lotus japonicus
1	G727	572293	1.20E-31	Oryza sativa
1	G727	7218448	7.70E-30	Sorghum bicolor
1	G727	9291284	1.80E-27	Lycopersicon hirsutum
1	G727	8901641	5.10E-27	Hordeum vulgare
1	G727	8380453	6.60E-24	Gossypium arboreum
1	G727	9962201	2.10E-12	Cryptomeria japonica
1	G727	8122498	3.10E-08	Lycopersicon esculentum
3	G732	5048074	5.60E-30	Gossypium hirsutum
3	G732	4384142	6.10E-30	Lycopersicon esculentum
3	G732	7623218	6.10E-30	Gossypium arboreum
3	G732	4457220	1.80E-29	Capsicum chinense
3	G732	7284989	4.50E-28	Glycine max
3	G732	9650827	1.20E-27	Petroselinum crispum
3	G732	7205618	2.20E-26	Medicago truncatula
3	G732	3854258	1.40E-22	Populus tremula x Populus tremuloides
5	G9	7643366	6.80E-56	Medicago truncatula
5	G9	8669779	4.20E-50	Glycine max
5	G9	8329389	1.50E-48	Mesembryanthemum crystallinum
5	G9	9851335	3.50E-42	Sorghum bicolor
5	G9	7412012	1.50E-41	Lycopersicon esculentum
5	G9	10450225	1.30E-38	Solanum tuberosum
5	G9	8902194	8.30E-36	Hordeum vulgare
5	G9	7722547	2.60E-33	Lotus japonicus
5	G9	9696857	1.90E-32	Triticum aestivum
5	G9	7324245	2.40E-32	Lycopersicon pennellii
7	G428	3327268	5.50E-65	Ipomoea nil
7	G428	4589883	1.20E-60	Nicotiana tabacum
7	G428	1814233	2.20E-56	Solanum tuberosum
7	G428	7581978	8.50E-56	Dendrobium grex Madame Thong-In
7	G428	4098241	1.50E-53	Lycopersicon esculentum
7	G428	4099825	1.30E-38	Picea mariana
7	G428	3462611	2.50E-38	Pisum sativum
7	G428	3928842	1.90E-37	Picea abies
7	G428	9699343	2.70E-35	Triticum aestivum
7	G428	1008878	4.80E-35	Zea mays
9	G869	10235055	1.00E-19	Glycine max
9	G869	2213784	1.60E-19	Lycopersicon esculentum
9	G869	3065894	9.20E-19	Nicotiana tabacum
9	G869	8570080	5.30E-18	Oryza sativa
9	G869	7560260	1.90E-17	Medicago truncatula
9	G869	9850452	9.30E-16	Sorghum bicolor
9	G869	9963144	1.10E-13	Cryptomeria japonica
9	G869	9660634	1.90E-13	Secale cereale
9	G869	9362061	3.40E-13	Triticum aestivum
9	G869	7788764	7.20E-13	Lotus japonicus
11	G1269	9565366	7.00E-37	Glycine max
11	G1269	5272360	8.10E-37	Lycopersicon esculentum
11	G1269	9119112	8.40E-28	Medicago truncatula
11	G1269	9852711	2.10E-22	Sorghum bicolor

Figure 3B

SEQ ID No.	GID	Genbank NID	P-value	Species
11	G1269	9255178	1.10E-18	Zea mays
11	G1269	10447957	8.60E-15	Solanum tuberosum
11	G1269	9435251	1.20E-09	Hordeum vulgare
11	G1269	3858030	3.20E-09	Populus balsamifera subsp. trichocarpa
11	G1269	9696112	3.80E-09	Triticum aestivum
11	G1269	8213273	4.90E-09	Oryza sativa
13	G1038	8748344	8.00E-37	Medicago truncatula
13	G1038	7283684	5.20E-36	Glycine max
13	G1038	7218448	8.80E-36	Sorghum bicolor
13	G1038	572293	3.30E-35	Oryza sativa
13	G1038	8901641	4.30E-28	Hordeum vulgare
13	G1038	9962201	2.20E-16	Cryptomeria japonica
13	G1038	7614196	6.50E-11	Lotus japonicus
13	G1038	9291272	0.00015	Lycopersicon hirsutum
13	G1038	8122498	0.0005	Lycopersicon esculentum
13	G1038	9883662	0.68	Triticum aestivum
15	G438	7209474	8.70E-204	Oryza sativa
15	G438	7209911	2.20E-142	Physcomitrella patens
15	G438	7571387	2.30E-80	Medicago truncatula
15	G438	8330425	3.00E-66	Mesembryanthemum crystallinum
15	G438	6531152	1.60E-64	Lycopersicon esculentum
15	G438	6726825	4.70E-61	Glycine max
15	G438	5269007	7.00E-54	Zea mays
15	G438	9253000	1.70E-47	Solanum tuberosum
15	G438	8967371	4.40E-46	Hordeum vulgare
15	G438	2963336	1.60E-34	Pinus taeda
17	G571	6288681	1.50E-70	Nicotiana tabacum
17	G571	297019	1.60E-68	Zea mays
17	G571	10423526	2.20E-61	Oryza sativa
17	G571	5926681	4.20E-61	Triticum aestivum
17	G571	4959969	1.90E-59	Lycopersicon esculentum
17	G571	1372965	1.20E-56	Vicia faba
17	G571	8098832	1.20E-46	Hordeum vulgare
17	G571	9566058	2.00E-43	Glycine max
17	G571	765198	1.50E-41	Solanum tuberosum
17	G571	19679	3.80E-41	Nicotiana sp.
19	G748	853689	7.00E-87	Cucurbita maxima
19	G748	7242897	3.90E-59	Oryza sativa
19	G748	5888560	1.20E-45	Lycopersicon esculentum
19	G748	6341666	5.60E-38	Glycine max
19	G748	10700058	1.10E-36	Medicago truncatula
19	G748	7535776	5.00E-33	Sorghum bicolor
19	G748	9419494	2.10E-31	Hordeum vulgare
19	G748	9410157	1.00E-28	Triticum aestivum
19	G748	3929324	4.30E-25	Dendrobium grex Madame Thong-IN
19	G748	10449922	2.30E-23	Solanum tuberosum
21	G431	7340349	9.90E-177	Brassica oleracea
21	G431	3462611	1.20E-112	Pisum sativum
21	G431	310568	1.50E-112	Glycine max
21	G431	2251078	1.90E-107	Nicotiana tabacum
21	G431	4098239	1.20E-104	Lycopersicon esculentum
21	G431	1008878	4.90E-62	Zea mays
21	G431	6942299	7.90E-62	Triticum aestivum

Figure 3C

SEQ ID No.	GID	Genbank NID	P-value	Species
21	G431	3327239	1.90E-61	Oryza sativa
21	G431	3928842	1.60E-59	Picea abies
21	G431	2522483	2.30E-59	Hordeum vulgare
23	G187	9304207	2.10E-35	Sorghum bicolor
23	G187	9444636	3.20E-34	Triticum aestivum
23	G187	5058292	3.60E-34	Glycine max
23	G187	7721184	2.40E-32	Lotus japonicus
23	G187	7562279	1.20E-31	Medicago truncatula
23	G187	8105974	3.00E-29	Lycopersicon esculentum
23	G187	9049477	1.60E-27	Oryza sativa
23	G187	9187621	1.60E-23	Solanum tuberosum
23	G187	5268376	5.60E-23	Zea mays
23	G187	4894964	1.70E-22	Avena sativa
25	G470	6917173	4.80E-78	Lycopersicon pennellii
25	G470	8827792	8.50E-70	Glycine max
25	G470	5272309	7.40E-69	Lycopersicon esculentum
25	G470	7563870	6.70E-68	Medicago truncatula
25	G470	5296108	5.50E-65	Zea mays
25	G470	7339690	7.40E-57	Oryza sativa
25	G470	5047367	1.30E-51	Gossypium hirsutum
25	G470	9856054	9.70E-50	Sorghum bicolor
25	G470	3857884	1.10E-38	Populus balsamifera subsp. trichocarpa
25	G470	8174666	6.40E-37	Hordeum vulgare
27	G615	5566284	2.00E-28	Linaria vulgaris
27	G615	6358617	3.20E-27	Antirrhinum graniticum
27	G615	6358613	1.40E-26	Antirrhinum majus subsp. cirrhigerum
27	G615	6358545	8.60E-26	Digitalis purpurea
27	G615	6358538	1.40E-25	Antirrhinum braun-blanquetii
27	G615	6358541	1.40E-25	Misopates orontium
27	G615	6358542	1.40E-25	Antirrhinum molle
27	G615	6358573	1.40E-25	Misopates calycinum
27	G615	6358546	1.80E-25	Antirrhinum siculum
27	G615	2826867	2.70E-25	Antirrhinum majus
29	G1073	7238733	2.70E-55	Medicago truncatula
29	G1073	10843924	1.50E-44	Glycine max
29	G1073	7615218	2.00E-42	Lotus japonicus
29	G1073	7333102	3.40E-34	Lycopersicon esculentum
29	G1073	9689692	8.60E-28	Pinus taeda
29	G1073	9445090	4.30E-25	Triticum aestivum
29	G1073	9252370	2.80E-24	Solanum tuberosum
29	G1073	5042437	5.80E-21	Oryza sativa
29	G1073	7536402	6.70E-20	Sorghum bicolor
29	G1073	9662742	2.70E-19	Secale cereale
31	G1493	7614196	2.20E-50	Lotus japonicus
31	G1493	9986889	6.10E-48	Glycine max
31	G1493	8748344	2.20E-38	Medicago truncatula
31	G1493	572293	1.70E-37	Oryza sativa
31	G1493	7218448	5.70E-33	Sorghum bicolor
31	G1493	9291284	9.70E-32	Lycopersicon hirsutum
31	G1493	8380453	1.60E-30	Gossypium arboreum
31	G1493	8901641	1.70E-30	Hordeum vulgare
31	G1493	9962201	6.90E-17	Cryptomeria japonica
31	G1493	8122498	1.50E-08	Lycopersicon esculentum

Figure 3D

SEQ ID No.	GID	Genbank NID	P-value	Species
33	G993	7643366	1.20E-58	Medicago truncatula
33	G993	8329389	1.00E-49	Mesembryanthemum crystallinum
33	G993	8669779	6.10E-49	Glycine max
33	G993	9851335	6.30E-43	Sorghum bicolor
33	G993	4384549	5.20E-40	Lycopersicon esculentum
33	G993	10450225	3.70E-39	Solanum tuberosum
33	G993	8902194	2.50E-34	Hordeum vulgare
33	G993	7719409	1.30E-32	Lotus japonicus
33	G993	8749037	5.20E-32	Citrus x paradisi
33	G993	9247126	1.30E-30	Oryza sativa
35	G867	7643366	2.20E-57	Medicago truncatula
35	G867	8329389	1.10E-50	Mesembryanthemum crystallinum
35	G867	8669779	2.70E-46	Glycine max
35	G867	10450225	3.60E-41	Solanum tuberosum
35	G867	9851335	2.80E-40	Sorghum bicolor
35	G867	9430646	7.20E-40	Lycopersicon esculentum
35	G867	8902194	1.60E-34	Hordeum vulgare
35	G867	7722547	1.30E-33	Lotus japonicus
35	G867	7324245	3.90E-32	Lycopersicon pennellii
35	G867	8749037	1.40E-31	Citrus x paradisi
37	G1930	7643366	9.70E-57	Medicago truncatula
37	G1930	8329389	4.50E-47	Mesembryanthemum crystallinum
37	G1930	6069592	1.10E-46	Glycine max
37	G1930	10450225	6.50E-42	Solanum tuberosum
37	G1930	9430646	8.20E-39	Lycopersicon esculentum
37	G1930	9851335	1.80E-38	Sorghum bicolor
37	G1930	7722547	4.70E-34	Lotus japonicus
37	G1930	7324245	1.20E-32	Lycopersicon pennellii
37	G1930	8902194	3.00E-31	Hordeum vulgare
37	G1930	9697984	4.60E-29	Triticum aestivum
39	G1594	3327268	2.60E-74	Ipomoea nil
39	G1594	7581978	9.20E-62	Dendrobium grex Madame Thong-In
39	G1594	4887609	1.50E-47	Oryza sativa
39	G1594	1814233	4.00E-46	Solanum tuberosum
39	G1594	4589883	6.30E-43	Nicotiana tabacum
39	G1594	4098241	6.70E-43	Lycopersicon esculentum
39	G1594	3928842	2.00E-42	Picea abies
39	G1594	4099825	2.60E-42	Picea mariana
39	G1594	4240538	1.70E-41	Zea mays
39	G1594	1946219	1.90E-41	Malus domestica
41	G391	7209474	4.70E-194	Oryza sativa
41	G391	7209911	2.10E-145	Physcomitrella patens
41	G391	7560927	8.70E-67	Medicago truncatula
41	G391	10808354	1.50E-61	Solanum tuberosum
41	G391	5893826	7.00E-60	Lycopersicon esculentum
41	G391	8330425	8.60E-59	Mesembryanthemum crystallinum
41	G391	8284059	8.70E-57	Glycine max
41	G391	5269007	8.10E-46	Zea mays
41	G391	9419425	1.70E-43	Hordeum vulgare
41	G391	2963336	2.10E-37	Pinus taeda
43	G390	7209474	2.50E-166	Oryza sativa
43	G390	7209911	1.70E-149	Physcomitrella patens
43	G390	7560927	5.80E-81	Medicago truncatula



Figure 3E

SEQ ID No.	GID	Genbank NID	P-value	Species
43	G390	7409018	1.80E-68	Lycopersicon esculentum
43	G390	8071613	3.00E-60	Solanum tuberosum
43	G390	9466042	1.60E-59	Hordeum vulgare
43	G390	8284059	1.00E-57	Glycine max
43	G390	8330425	2.60E-44	Mesembryanthemum crystallinum
43	G390	5269007	4.60E-44	Zea mays
43	G390	2963336	4.90E-43	Pinus taeda
45	G1548	7209474	5.90E-169	Oryza sativa
45	G1548	7209911	3.30E-140	Physcomitrella patens
45	G1548	9253000	1.60E-76	Solanum tuberosum
45	G1548	9820423	1.40E-67	Glycine max
45	G1548	7570825	8.40E-67	Medicago truncatula
45	G1548	9456848	2.70E-55	Lycopersicon esculentum
45	G1548	9419425	1.40E-47	Hordeum vulgare
45	G1548	6626571	3.50E-46	Zea mays
45	G1548	8330425	4.20E-46	Mesembryanthemum crystallinum
45	G1548	3853847	2.70E-42	Populus tremula x Populus tremuloides

MBI0018 Sequence Listing.ST25  
SEQUENCE LISTING

```

<110> Riechmann, Jose Luis
      Reuber, Lynne
      Keddie, James
      Ratcliffe, Oliver
      Heard, Jacqueline
      Samaha, Raymond
      Yu, Guo-Liang
      Jiang, Cai-Zhong

<120> Plant Developmental Genes

<130> MBI-0018

<150> 60/164,132
<151> 1999-11-17

<150> 60/197,899
<151> 2000-04-17

<150> Plant Trait Modification III
<151> 2000-08-22

<160> 46

<170> PatentIn version 3.0

<210> 1
<211> 2007
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (43)..(1977)
<223> G727

<400> 1
cttcttctcc ttctctgata gttcgttttc tggacgagag ag atg gta aat ccg      54
                               Met Val Asn Pro
                               1

ggt cac gga aga gga ccc gat tcg ggt act gct gct ggt ggg tca aac      102
Gly His Gly Arg Gly Pro Asp Ser Gly Thr Ala Ala Gly Gly Ser Asn
5                               10                               15                               20

tcc gac ccg ttt cct gcg aat ctt cga gtt ctt gtc gtt gat gat gat      150
Ser Asp Pro Phe Pro Ala Asn Leu Arg Val Leu Val Val Asp Asp Asp
25                               30                               35

cca act tgt ctc atg atc tta gag agg atg ctt atg act tgt ctc tac      198
Pro Thr Cys Leu Met Ile Leu Glu Arg Met Leu Met Thr Cys Leu Tyr
40                               45                               50

aga gag cag aga gcg cat tgt ctc tgc ttc gga aga aca aag aat ggt      246
Arg Glu Gln Arg Ala His Cys Leu Cys Phe Gly Arg Thr Lys Asn Gly
55                               60                               65

ttt gat att gtc att agt gat gtt cat atg cct gac atg gat ggt ttc      294
Phe Asp Ile Val Ile Ser Asp Val His Met Pro Asp Met Asp Gly Phe
70                               75                               80

aag ctc ctt gaa cac gtt ggt tta gag atg gat tta cct gtt atc aat      342
Lys Leu Leu Glu His Val Gly Leu Glu Met Asp Leu Pro Val Ile Asn
85                               90                               95                               100

ctg aat gtt ttg aaa cct ttg gtt ata gtg atg tct gcg gat gat tcg      390
Leu Asn Val Leu Lys Pro Leu Val Ile Val Met Ser Ala Asp Asp Ser
105                               110                               115

aag agc gtt gtg ttg aaa gga gtg act cac ggt gca gtt gat tac ctc      438
Lys Ser Val Val Leu Lys Gly Val Thr His Gly Ala Val Asp Tyr Leu
120                               125                               130

```

## MBI0018 Sequence Listing.ST25

atc	aaa	ccg	gta	cgt	att	gag	gct	ttg	aag	aat	ata	tgg	caa	cat	gtg	486
Ile	Lys	Pro	Val	Arg	Ile	Glu	Ala	Leu	Lys	Asn	Ile	Trp	Gln	His	Val	
		135					140					145				
gtg	cgg	aag	aag	cgt	aac	gag	tgg	aat	gtt	tct	gaa	cat	tct	gga	gga	534
Val	Arg	Lys	Lys	Arg	Asn	Glu	Trp	Asn	Val	Ser	Glu	His	Ser	Gly	Gly	
		150				155					160					
agt	att	gaa	gat	act	ggc	ggt	gac	agg	gac	agg	cag	cag	cag	cat	agg	582
Ser	Ile	Glu	Asp	Thr	Gly	Gly	Asp	Arg	Asp	Arg	Gln	Gln	Gln	His	Arg	
		165			170				175						180	
gag	gat	gct	gat	aac	aac	tcg	tct	tca	gtt	aat	gaa	ggg	aac	ggg	agg	630
Glu	Asp	Ala	Asp	Asn	Asn	Ser	Ser	Ser	Val	Asn	Glu	Gly	Asn	Gly	Arg	
				185					190					195		
agc	tcg	agg	aag	cgg	aag	gaa	gag	gaa	gta	gat	gat	caa	ggg	gat	gat	678
Ser	Ser	Arg	Lys	Arg	Lys	Glu	Glu	Glu	Val	Asp	Asp	Gln	Gly	Asp	Asp	
			200					205					210			
aag	gaa	gac	tca	tcg	agt	tta	aag	aaa	cca	cgc	gtg	gtt	tgg	tct	gtt	726
Lys	Glu	Asp	Ser	Ser	Ser	Leu	Lys	Lys	Pro	Arg	Val	Val	Trp	Ser	Val	
		215					220					225				
gaa	ttg	cat	cag	cag	ttt	gtt	gct	gct	gtg	aat	cag	cta	ggc	gtt	gac	774
Glu	Leu	His	Gln	Gln	Phe	Val	Ala	Ala	Val	Asn	Gln	Leu	Gly	Val	Asp	
		230				235					240					
agt	gag	tta	aaa	act	tgc	ttg	ctt	atg	cat	ttg	tgt	gtg	tcg	att	ggc	822
Ser	Glu	Leu	Lys	Thr	Cys	Leu	Leu	Met	His	Leu	Cys	Val	Ser	Ile	Gly	
		245			250					255					260	
aac	att	gtg	gaa	ttc	cag	aag	tat	cgg	ata	tat	ctg	aga	cgg	ctt	gga	870
Asn	Ile	Val	Glu	Phe	Gln	Lys	Tyr	Arg	Ile	Tyr	Leu	Arg	Arg	Leu	Gly	
				265					270					275		
gga	gta	tcg	caa	cac	caa	gga	aat	atg	aac	cat	tcg	ttt	atg	act	ggc	918
Gly	Val	Ser	Gln	His	Gln	Gly	Asn	Met	Asn	His	Ser	Phe	Met	Thr	Gly	
			280					285					290			
caa	gat	cag	agt	ttt	gga	cct	ctt	tct	tcg	ttg	aat	gga	ttt	gat	ctt	966
Gln	Asp	Gln	Ser	Phe	Gly	Pro	Leu	Ser	Ser	Leu	Asn	Gly	Phe	Asp	Leu	
		295				300						305				
caa	tct	tta	gct	gtt	act	ggt	cag	ctc	cct	cct	cag	agc	ctt	gca	cag	1014
Gln	Ser	Leu	Ala	Val	Thr	Gly	Gln	Leu	Pro	Pro	Gln	Ser	Leu	Ala	Gln	
		310				315					320					
ctt	caa	gca	gct	ggt	ctt	ggc	cgg	cct	aca	ctc	gct	aaa	cca	ggg	atg	1062
Leu	Gln	Ala	Ala	Gly	Leu	Gly	Arg	Pro	Thr	Leu	Ala	Lys	Pro	Gly	Met	
		325			330					335					340	
tcg	gtt	tct	ccc	ctt	gta	gat	cag	aga	agc	atc	ttc	aac	ttt	gaa	aac	1110
Ser	Val	Ser	Pro	Leu	Val	Asp	Gln	Arg	Ser	Ile	Phe	Asn	Phe	Glu	Asn	
			345						350					355		
cca	aaa	ata	aga	ttt	gga	gac	gga	cat	ggt	cag	acg	atg	aac	aat	gga	1158
Pro	Lys	Ile	Arg	Phe	Gly	Asp	Gly	His	Gly	Gln	Thr	Met	Asn	Asn	Gly	
			360					365					370			
aat	ttg	ctt	cat	ggt	gtc	cca	acg	ggt	agt	cac	atg	cgt	ctg	cgt	cct	1206
Asn	Leu	Leu	His	Gly	Val	Pro	Thr	Gly	Ser	His	Met	Arg	Leu	Arg	Pro	
		375					380					385				
gga	cag	aat	gtt	cag	agc	agc	gga	atg	atg	ttg	cca	gta	gca	gac	cag	1254
Gly	Gln	Asn	Val	Gln	Ser	Ser	Gly	Met	Met	Leu	Pro	Val	Ala	Asp	Gln	
		390				395				400						
cta	cct	cga	gga	gga	cca	tcg	atg	cta	cca	tcc	ctc	ggg	caa	cag	ccg	1302
Leu	Pro	Arg	Gly	Gly	Pro	Ser	Met	Leu	Pro	Ser	Leu	Gly	Gln	Gln	Pro	
		405			410					415					420	
ata	ttg	tca	agc	agc	gtt	tca	aga	aga	agc	gat	ctc	act	ggt	gcg	ctg	1350
Ile	Leu	Ser	Ser	Ser	Val	Ser	Arg	Arg	Ser	Asp	Leu	Thr	Gly	Ala	Leu	

## MBI0018 Sequence Listing.ST25

425	430	435	
gcg gtt aga aac agt atc ccc gag acc aac agc aga gtg tta cca act Ala Val Arg Asn Ser Ile Pro Glu Thr Asn Ser Arg Val Leu Pro Thr 440 445 450			1398
act cac tcg gtc ttc aat aac ttc ccc gcg gat cta cct cgc agc agc Thr His Ser Val Phe Asn Asn Phe Pro Ala Asp Leu Pro Arg Ser Ser 455 460 465			1446
ttc ccg ttg gca agt gcc cca ggg att tca gtt cca gta tca gtt tct Phe Pro Leu Ala Ser Ala Pro Gly Ile Ser Val Pro Val Ser Val Ser 470 475 480			1494
tac caa gaa gag gtc aac agc tcg gat gca aaa gga ggt tca tca gct Tyr Gln Glu Glu Val Asn Ser Ser Asp Ala Lys Gly Gly Ser Ser Ala 485 490 500			1542
gct act gct gga ttt ggt aac cca agc tac gac ata ttt aac gat ttt Ala Thr Ala Gly Phe Gly Asn Pro Ser Tyr Asp Ile Phe Asn Asp Phe 505 510 515			1590
ccg cag cac caa cag cac aac aag aac atc agc aat aaa cta aac gat Pro Gln His Gln Gln His Asn Lys Asn Ile Ser Asn Lys Leu Asn Asp 520 525 530			1638
tgg gat ctg cgg aat atg gga ttg gtc ttc agt tcc aat cag gac gca Trp Asp Leu Arg Asn Met Gly Leu Val Phe Ser Ser Asn Gln Asp Ala 535 540 545			1686
gca act gca acc gca acc gca gca ttt tcc act tcg gaa gca tac tct Ala Thr Ala Thr Ala Thr Ala Ala Phe Ser Thr Ser Glu Ala Tyr Ser 550 555 560			1734
tcg tct tct acg cag aga aaa aga cgg gaa acg gac gca aca gtt gtg Ser Ser Ser Thr Gln Arg Lys Arg Arg Glu Thr Asp Ala Thr Val Val 565 570 575 580			1782
ggt gag cat ggg cag aac ctg cag tca ccg agc cgg aat ctg tat cat Gly Glu His Gly Gln Asn Leu Gln Ser Pro Ser Arg Asn Leu Tyr His 585 590 595			1830
ctg aac cac gtt ttt atg gac ggt ggt tca gtc aga gtg aag tca gaa Leu Asn His Val Phe Met Asp Gly Gly Ser Val Arg Val Lys Ser Glu 600 605 610			1878
aga gtg gcg gag aca gtg act tgt cct cca gca aat aca ttg ttt cac Arg Val Ala Glu Thr Val Thr Cys Pro Pro Ala Asn Thr Leu Phe His 615 620 625			1926
gag cag tat aat caa gaa gat ctg atg agc gca ttt ctc aaa cag gtt Glu Gln Tyr Asn Gln Glu Asp Leu Met Ser Ala Phe Leu Lys Gln Val 630 635 640			1974
tga ttattactcg aatacagtgc actctaaaaac			2007
<210> 2			
<211> 644			
<212> PRT			
<213> Arabidopsis thaliana			
<400> 2			
Met Val Asn Pro Gly His Gly Arg Gly Pro Asp Ser Gly Thr Ala Ala 1 5 10 15			
Gly Gly Ser Asn Ser Asp Pro Phe Pro Ala Asn Leu Arg Val Leu Val 20 25 30			
Val Asp Asp Asp Pro Thr Cys Leu Met Ile Leu Glu Arg Met Leu Met 35 40 45			

## MBI0018 Sequence Listing.ST25

Thr Cys Leu Tyr Arg Glu Gln Arg Ala His Cys Leu Cys Phe Gly Arg  
 50 55 60  
 Thr Lys Asn Gly Phe Asp Ile Val Ile Ser Asp Val His Met Pro Asp  
 65 70 75 80  
 Met Asp Gly Phe Lys Leu Leu Glu His Val Gly Leu Glu Met Asp Leu  
 85 90 95  
 Pro Val Ile Asn Leu Asn Val Leu Lys Pro Leu Val Ile Val Met Ser  
 100 105 110  
 Ala Asp Asp Ser Lys Ser Val Val Leu Lys Gly Val Thr His Gly Ala  
 115 120 125  
 Val Asp Tyr Leu Ile Lys Pro Val Arg Ile Glu Ala Leu Lys Asn Ile  
 130 135 140  
 Trp Gln His Val Val Arg Lys Lys Arg Asn Glu Trp Asn Val Ser Glu  
 145 150 155 160  
 His Ser Gly Gly Ser Ile Glu Asp Thr Gly Gly Asp Arg Asp Arg Gln  
 165 170 175  
 Gln Gln His Arg Glu Asp Ala Asp Asn Asn Ser Ser Ser Val Asn Glu  
 180 185 190  
 Gly Asn Gly Arg Ser Ser Arg Lys Arg Lys Glu Glu Glu Val Asp Asp  
 195 200 205  
 Gln Gly Asp Asp Lys Glu Asp Ser Ser Ser Leu Lys Lys Pro Arg Val  
 210 215 220  
 Val Trp Ser Val Glu Leu His Gln Gln Phe Val Ala Ala Val Asn Gln  
 225 230 235 240  
 Leu Gly Val Asp Ser Glu Leu Lys Thr Cys Leu Leu Met His Leu Cys  
 245 250 255  
 Val Ser Ile Gly Asn Ile Val Glu Phe Gln Lys Tyr Arg Ile Tyr Leu  
 260 265 270  
 Arg Arg Leu Gly Gly Val Ser Gln His Gln Gly Asn Met Asn His Ser  
 275 280 285  
 Phe Met Thr Gly Gln Asp Gln Ser Phe Gly Pro Leu Ser Ser Leu Asn  
 290 295 300  
 Gly Phe Asp Leu Gln Ser Leu Ala Val Thr Gly Gln Leu Pro Pro Gln  
 305 310 315 320  
 Ser Leu Ala Gln Leu Gln Ala Ala Gly Leu Gly Arg Pro Thr Leu Ala  
 325 330 335  
 Lys Pro Gly Met Ser Val Ser Pro Leu Val Asp Gln Arg Ser Ile Phe  
 340 345 350

## MBI0018 Sequence Listing.ST25

Asn Phe Glu Asn Pro Lys Ile Arg Phe Gly Asp Gly His Gly Gln Thr  
 355 360 365  
 Met Asn Asn Gly Asn Leu Leu His Gly Val Pro Thr Gly Ser His Met  
 370 375 380  
 Arg Leu Arg Pro Gly Gln Asn Val Gln Ser Ser Gly Met Met Leu Pro  
 385 390 395 400  
 Val Ala Asp Gln Leu Pro Arg Gly Gly Pro Ser Met Leu Pro Ser Leu  
 405 410 415  
 Gly Gln Gln Pro Ile Leu Ser Ser Ser Val Ser Arg Arg Ser Asp Leu  
 420 425 430  
 Thr Gly Ala Leu Ala Val Arg Asn Ser Ile Pro Glu Thr Asn Ser Arg  
 435 440 445  
 Val Leu Pro Thr Thr His Ser Val Phe Asn Asn Phe Pro Ala Asp Leu  
 450 455 460  
 Pro Arg Ser Ser Phe Pro Leu Ala Ser Ala Pro Gly Ile Ser Val Pro  
 465 470 475 480  
 Val Ser Val Ser Tyr Gln Glu Glu Val Asn Ser Ser Asp Ala Lys Gly  
 485 490 495  
 Gly Ser Ser Ala Ala Thr Ala Gly Phe Gly Asn Pro Ser Tyr Asp Ile  
 500 505 510  
 Phe Asn Asp Phe Pro Gln His Gln Gln His Asn Lys Asn Ile Ser Asn  
 515 520 525  
 Lys Leu Asn Asp Trp Asp Leu Arg Asn Met Gly Leu Val Phe Ser Ser  
 530 535 540  
 Asn Gln Asp Ala Ala Thr Ala Thr Ala Thr Ala Phe Ser Thr Ser  
 545 550 555 560  
 Glu Ala Tyr Ser Ser Ser Ser Thr Gln Arg Lys Arg Arg Glu Thr Asp  
 565 570 575  
 Ala Thr Val Val Gly Glu His Gly Gln Asn Leu Gln Ser Pro Ser Arg  
 580 585 590  
 Asn Leu Tyr His Leu Asn His Val Phe Met Asp Gly Gly Ser Val Arg  
 595 600 605  
 Val Lys Ser Glu Arg Val Ala Glu Thr Val Thr Cys Pro Pro Ala Asn  
 610 615 620  
 Thr Leu Phe His Glu Gln Tyr Asn Gln Glu Asp Leu Met Ser Ala Phe  
 625 630 635 640  
 Leu Lys Gln Val

## MBI0018 Sequence Listing.ST25

```

<210> 3
<211> 834
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (73)..(588)
<223> G732

<400> 3
aaaaaaacca aacataaaac ataaaactct gtcctttttt tgtcttcttg taacttttct 60
tgttaaaaat ca atg gcg tca tct agc agc aca tac cgg agc tca agc tct 111
          Met Ala Ser Ser Ser Ser Thr Tyr Arg Ser Ser Ser Ser
          1                    5                    10

tcc gac ggt ggt aat aat aac ccg tcg gac tcc gtc gtc acc gtc gac 159
Ser Asp Gly Gly Asn Asn Asn Pro Ser Asp Ser Val Val Thr Val Asp
15                    20                    25

gaa cga aaa cgt aaa aga atg tta tcg aac aga gaa tct gca cgt agg 207
Glu Arg Lys Arg Lys Arg Met Leu Ser Asn Arg Glu Ser Ala Arg Arg
30                    35                    40                    45

tca agg atg cgt aaa cag aaa cac gtt gat gat cta acg gct cag atc 255
Ser Arg Met Arg Lys Gln Lys His Val Asp Asp Leu Thr Ala Gln Ile
50                    55                    60

aat cag cta tca aac gac aac cgt cag atc ttg aac agc ctc acc gta 303
Asn Gln Leu Ser Asn Asp Asn Arg Gln Ile Leu Asn Ser Leu Thr Val
65                    70                    75

aca tct cag ctt tac atg aag atc caa gcc gag aac tct gtt ctc acc 351
Thr Ser Gln Leu Tyr Met Lys Ile Gln Ala Glu Asn Ser Val Leu Thr
80                    85                    90

gct cag atg gag gag ctt agc acc aga ctc caa tct ctc aac gag atc 399
Ala Gln Met Glu Glu Leu Ser Thr Arg Leu Gln Ser Leu Asn Glu Ile
95                    100                    105

gtt gat ctt gtt caa tcc aac ggt gca gga ttt ggt gtt gac cag atc 447
Val Asp Leu Val Gln Ser Asn Gly Ala Gly Phe Gly Val Asp Gln Ile
110                    115                    120                    125

gac ggc tgt ggt ttt gat gat cgt acg gtt ggg atc gac gga tat tac 495
Asp Gly Cys Gly Phe Asp Asp Arg Thr Val Gly Ile Asp Gly Tyr Tyr
130                    135                    140

gat gat atg aat atg atg agt aat gtt aat cat tgg ggt ggt tcg gtt 543
Asp Asp Met Asn Met Met Ser Asn Val Asn His Trp Gly Gly Ser Val
145                    150                    155

tac act aac caa ccc att atg gct aat gat atc aat atg tat tga 588
Tyr Thr Asn Gln Pro Ile Met Ala Asn Asp Ile Asn Met Tyr
160                    165                    170

ttaataaaaat taattaaaat aattagatgc cccttttttg tcttttttatt ttaaaattta 648
gcccatTTTTg gtgttttttg gttggtgtga tgatgtaatt atagtacatg catctttgat 708
tggttggaag gataaatata aactttatat atatattggg gcatatatat atgagttgta 768
ctttgcatgt attggtgtgt gttttgttat aattatatga ttatatatgt ttatgttaaa 828
aaaaaa 834

<210> 4
<211> 171
<212> PRT

```

## MBI0018 Sequence Listing.ST25

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 4

Met Ala Ser Ser Ser Ser Thr Tyr Arg Ser Ser Ser Ser Ser Asp Gly  
1 5 10 15

Gly Asn Asn Asn Pro Ser Asp Ser Val Val Thr Val Asp Glu Arg Lys  
20 25 30

Arg Lys Arg Met Leu Ser Asn Arg Glu Ser Ala Arg Arg Ser Arg Met  
35 40 45

Arg Lys Gln Lys His Val Asp Asp Leu Thr Ala Gln Ile Asn Gln Leu  
50 55 60

Ser Asn Asp Asn Arg Gln Ile Leu Asn Ser Leu Thr Val Thr Ser Gln  
65 70 75 80

Leu Tyr Met Lys Ile Gln Ala Glu Asn Ser Val Leu Thr Ala Gln Met  
85 90 95

Glu Glu Leu Ser Thr Arg Leu Gln Ser Leu Asn Glu Ile Val Asp Leu  
100 105 110

Val Gln Ser Asn Gly Ala Gly Phe Gly Val Asp Gln Ile Asp Gly Cys  
115 120 125

Gly Phe Asp Asp Arg Thr Val Gly Ile Asp Gly Tyr Tyr Asp Asp Met  
130 135 140

Asn Met Met Ser Asn Val Asn His Trp Gly Gly Ser Val Tyr Thr Asn  
145 150 155 160

Gln Pro Ile Met Ala Asn Asp Ile Asn Met Tyr  
165 170

&lt;210&gt; 5

&lt;211&gt; 1246

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (81)..(1139)

&lt;223&gt; G9

&lt;400&gt; 5

gtgtttcttc tttctgctaa aaggttataa tttttgtttc ttggtttggt gagaatcttc 60

aagaaactga aacaaagaaa atg gat tct agt tgc ata gac gag ata agt tcc 113  
Met Asp Ser Ser Cys Ile Asp Glu Ile Ser Ser  
1 5 10

tcc act tca gaa tct ttc tcc gcc acc acc gcc aag aag ctc tct cct 161  
Ser Thr Ser Glu Ser Phe Ser Ala Thr Thr Ala Lys Lys Leu Ser Pro  
15 20 25

cct ccc gcg gcg gcg tta cgc ctc tac cgg atg gga agc ggc ggg agc 209  
Pro Pro Ala Glu Ala Ala Leu Arg Leu Tyr Arg Met Gly Ser Gly Gly Ser  
30 35 40

agc gtc gtg ttg gat ccc gag aac ggc cta gag acg gag tca cga aag 257



MBI0018 Sequence Listing.ST25																
Ser	Val	Val	Leu	Asp	Pro	Glu	Asn	Gly	Leu	Glu	Thr	Glu	Ser	Arg	Lys	
45						50					55					
cta	cca	tct	tca	aaa	tac	aaa	ggt	ggt	ggt	cct	cag	cct	aac	gga	aga	305
Leu	Pro	Ser	Ser	Lys	Tyr	Lys	Gly	Val	Val	Pro	Gln	Pro	Asn	Gly	Arg	
60					65					70					75	
tgg	gga	gct	cag	atc	tac	gag	aag	cac	caa	cga	gta	tgg	ctc	ggg	act	353
Trp	Gly	Ala	Gln	Ile	Tyr	Glu	Lys	His	Gln	Arg	Val	Trp	Leu	Gly	Thr	
				80					85					90		
ttc	aac	gag	caa	gaa	gaa	gct	gct	cgt	tcc	tac	gac	atc	gca	gct	tgt	401
Phe	Asn	Glu	Gln	Glu	Glu	Ala	Ala	Arg	Ser	Tyr	Asp	Ile	Ala	Ala	Cys	
			95					100					105			
aga	ttc	cgt	ggc	cgc	gac	gcc	gtc	gtc	aac	ttc	aag	aac	ggt	ctg	gaa	449
Arg	Phe	Arg	Gly	Arg	Asp	Ala	Val	Val	Asn	Phe	Lys	Asn	Val	Leu	Glu	
		110					115					120				
gac	ggc	gat	tta	gct	ttt	ctt	gaa	gct	cac	tca	aag	gcc	gag	atc	gtc	497
Asp	Gly	Asp	Leu	Ala	Phe		Glu	Ala	His	Ser	Lys	Ala	Glu	Ile	Val	
	125					130					135					
gac	atg	ttg	aga	aaa	cac	act	tac	gcc	gac	gag	ctt	gaa	cag	aac	aat	545
Asp	Met	Leu	Arg	Lys	His	Thr	Tyr	Ala	Asp	Glu	Leu	Glu	Gln	Asn	Asn	
					145					150					155	
aaa	cgg	cag	ttg	ttt	ctc	tcc	gtc	gac	gct	aac	gga	aaa	cgt	aac	gga	593
Lys	Arg	Gln	Leu	Phe	Leu	Ser	Val	Asp	Ala	Asn	Gly	Lys	Arg	Asn	Gly	
				160					165					170		
tcg	agt	act	act	caa	aac	gac	aaa	ggt	tta	aag	acg	tgt	gaa	ggt	ctt	641
Ser	Ser	Thr	Thr	Gln	Asn	Asp	Lys	Val	Leu	Lys	Thr	Cys	Glu	Val	Leu	
			175					180					185			
ttc	gag	aag	gct	ggt	aca	cct	agc	gac	ggt	ggg	aag	cta	aac	cgt	ctc	689
Phe	Glu	Lys	Ala	Val	Thr	Pro	Ser	Asp	Val	Gly	Lys	Leu	Asn	Arg	Leu	
		190					195					200				
gtg	ata	cct	aaa	caa	cac	gcc	gag	aaa	cac	ttt	ccg	tta	ccg	tca	ccg	737
Val	Ile	Pro	Lys	Gln	His	Ala	Glu	Lys	His	Phe	Pro	Leu	Pro	Ser	Pro	
	205					210					215					
tca	ccg	gca	gtg	act	aaa	gga	ggt	ttg	atc	aac	ttc	gaa	gac	ggt	aac	785
Ser	Pro	Ala	Val	Thr	Lys	Gly	Val	Leu	Ile	Asn	Phe	Glu	Asp	Val	Asn	
					225					230				235		
ggt	aaa	gtg	tgg	agg	ttc	cgt	tac	tca	tac	tgg	aac	agt	agt	caa	agt	833
Gly	Lys	Val	Trp	Arg	Phe	Arg	Tyr	Ser	Tyr	Trp	Asn	Ser	Ser	Gln	Ser	
				240					245					250		
tac	gtg	ttg	acc	aag	gga	tgg	agt	cga	ttc	gtc	aag	gag	aag	aat	ctt	881
Tyr	Val	Leu	Thr	Lys	Gly	Trp	Ser	Arg	Phe	Val	Lys	Glu	Lys	Asn	Leu	
			255					260					265			
cga	gcc	ggt	gat	ggt	ggt	act	ttc	gag	aga	tcg	acc	gga	cta	gag	cgg	929
Arg	Ala	Gly	Asp	Val	Val	Thr	Phe	Glu	Arg	Ser	Thr	Gly	Leu	Glu	Arg	
		270					275					280				
cag	tta	tat	att	gat	tgg	aaa	ggt	cgg	tct	ggt	ccg	aga	gaa	aac	ccg	977
Gln	Leu	Tyr	Ile	Asp	Trp	Lys	Val	Arg	Ser	Gly	Pro	Arg	Glu	Asn	Pro	
		285				290					295					
ggt	cag	gtg	gtg	ggt	cgg	ctt	ttc	gga	ggt	gat	atc	ttt	aat	gtg	acc	1025
Val	Gln	Val	Val	Val	Arg	Leu	Phe	Gly	Val	Asp	Ile	Phe	Asn	Val	Thr	
					305					310				315		
acc	gtg	aag	cca	aac	gac	gtc	gtg	gcc	ggt	tgc	ggt	gga	aag	aga	tct	1073
Thr	Val	Lys	Pro	Asn	Asp	Val	Val	Ala	Val	Cys	Gly	Gly	Lys	Arg	Ser	
				320					325					330		
cga	gat	ggt	gat	gat	atg	ttt	gcg	tta	cgg	tgt	tcc	aag	aag	cag	gcg	1121
Arg	Asp	Val	Asp	Asp	Met	Phe	Ala	Leu	Arg	Cys	Ser	Lys	Lys	Gln	Ala	
				335				340					345			

MBI0018 Sequence Listing.ST25

ata atc aat gct ttg tga catatttcct tttccgattt tatgctttcg 1169  
 ile ile asn ala leu  
 350

ttttttaatt tttttttttg tcaagttgtg tagggttgga ttcattgctag gttgtattta 1229

ggaaaagaga taagacc 1246

<210> 6  
 <211> 352  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 6

Met Asp Ser Ser Cys Ile Asp Glu Ile Ser Ser Ser Thr Ser Glu Ser  
 1 5 10 15

Phe Ser Ala Thr Thr Ala Lys Lys Leu Ser Pro Pro Pro Ala Ala Ala  
 20 25 30

Leu Arg Leu Tyr Arg Met Gly Ser Gly Gly Ser Ser Val Val Leu Asp  
 35 40 45

Pro Glu Asn Gly Leu Glu Thr Glu Ser Arg Lys Leu Pro Ser Ser Lys  
 50 55 60

Tyr Lys Gly Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln Ile  
 65 70 75 80

Tyr Glu Lys His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Gln Glu  
 85 90 95

Glu Ala Ala Arg Ser Tyr Asp Ile Ala Ala Cys Arg Phe Arg Gly Arg  
 100 105 110

Asp Ala Val Val Asn Phe Lys Asn Val Leu Glu Asp Gly Asp Leu Ala  
 115 120 125

Phe Leu Glu Ala His Ser Lys Ala Glu Ile Val Asp Met Leu Arg Lys  
 130 135 140

His Thr Tyr Ala Asp Glu Leu Glu Gln Asn Asn Lys Arg Gln Leu Phe  
 145 150 155 160

Leu Ser Val Asp Ala Asn Gly Lys Arg Asn Gly Ser Ser Thr Thr Gln  
 165 170 175

Asn Asp Lys Val Leu Lys Thr Cys Glu Val Leu Phe Glu Lys Ala Val  
 180 185 190

Thr Pro Ser Asp Val Gly Lys Leu Asn Arg Leu Val Ile Pro Lys Gln  
 195 200 205

His Ala Glu Lys His Phe Pro Leu Pro Ser Pro Ser Pro Ala Val Thr  
 210 215 220

Lys Gly Val Leu Ile Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg  
 225 230 235 240

## MBI0018 Sequence Listing.ST25

Phe Arg Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys  
245 250 255

Gly Trp Ser Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val  
260 265 270

Val Thr Phe Glu Arg Ser Thr Gly Leu Glu Arg Gln Leu Tyr Ile Asp  
275 280 285

Trp Lys Val Arg Ser Gly Pro Arg Glu Asn Pro Val Gln Val Val Val  
290 295 300

Arg Leu Phe Gly Val Asp Ile Phe Asn Val Thr Thr Val Lys Pro Asn  
305 310 315 320

Asp Val Val Ala Val Cys Gly Gly Lys Arg Ser Arg Asp Val Asp Asp  
325 330 335

Met Phe Ala Leu Arg Cys Ser Lys Lys Gln Ala Ile Ile Asn Ala Leu  
340 345 350

<210> 7  
<211> 1379  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (97)..(1032)  
<223> G428

<400> 7  
ttacttttgt gtttcttcat attcttcaga agcaagcaca aggctaggga tcgaagaagc 60

ggcgatcact gatcgtatct cactacgatac acatta atg gat aga atg tgt ggt 114  
Met Asp Arg Met Cys Gly  
1 5

ttc cgc tcg acg gaa gac tat tcg gag aaa gcg acg ttg atg atg ccg 162  
Phe Arg Ser Thr Glu Asp Tyr Ser Glu Lys Ala Thr Leu Met Met Pro  
10 15 20

tcc gat tat cag tct ttg att tgt tca acc acc gga gac aat caa aga 210  
Ser Asp Tyr Gln Ser Leu Ile Cys Ser Thr Thr Gly Asp Asn Gln Arg  
25 30 35

ctg ttt gga tcc gac gaa ctc gct acc gct ttg tcc tcg gag ttg ctt 258  
Leu Phe Gly Ser Asp Glu Leu Ala Thr Ala Leu Ser Ser Glu Leu Leu  
40 45 50

ccg cgt att cga aaa gct gag gat aat ttc tct ctt agt gtc atc aaa 306  
Pro Arg Ile Arg Lys Ala Glu Asp Asn Phe Ser Leu Ser Val Ile Lys  
55 60 65 70

tcc aaa atc gct tct cat cct ttg tat cct cgc tta ctc caa acc tac 354  
Ser Lys Ile Ala Ser His Pro Leu Tyr Pro Arg Leu Leu Gln Thr Tyr  
75 80 85

atc gat tgc caa aag gtg gga gcg cct atg gaa ata gcg tgt ata ttg 402  
Ile Asp Cys Gln Lys Val Gly Ala Pro Met Glu Ile Ala Cys Ile Leu  
90 95 100

gaa gag att cag cga gag aac cat gtg tac aag aga gat gtt gct cca 450  
Glu Glu Ile Gln Arg Glu Asn His Val Tyr Lys Arg Asp Val Ala Pro  
105 110 115

MBI0018 Sequence Listing.ST25

tta tct tgc ttt gga gct gat cct gag ctt gat gaa ttc atg gaa acc	498
Leu Ser Cys Phe Gly Ala Asp Pro Glu Leu Asp Glu Phe Met Glu Thr	
120 125 130	
tac tgt gat ata ttg gtt aaa tac aaa acc gat ctt gcg agg ccg ttc	546
Tyr Cys Asp Ile Leu Val Lys Tyr Lys Thr Asp Leu Ala Arg Pro Phe	
135 140 145 150	
gac gag gct aca act ttc ata aac aag att gaa atg cag ctt cag aac	594
Asp Glu Ala Thr Thr Phe Ile Asn Lys Ile Glu Met Gln Leu Gln Asn	
155 160 165	
ttg tgc act ggt cca gcg tct gct aca gct ctt tca gat gat ggt gcg	642
Leu Cys Thr Gly Pro Ala Ser Ala Thr Ala Leu Ser Asp Asp Gly Ala	
170 175 180	
gtt tca tct gac gag gaa ctg aga gaa gat gat gac ata gca gcg gat	690
Val Ser Ser Asp Glu Glu Leu Arg Glu Asp Asp Asp Ile Ala Ala Asp	
185 190 195	
gac agc caa caa aga agc aat gac cgc gat ctg aag gac cag cta cta	738
Asp Ser Gln Gln Arg Ser Asn Asp Arg Asp Leu Lys Asp Gln Leu Leu	
200 205 210	
cgc aaa ttt ggt agc cat atc agt tca ttg aaa ctc gag ttc tct aaa	786
Arg Lys Phe Gly Ser His Ile Ser Ser Leu Lys Leu Glu Phe Ser Lys	
215 220 225 230	
aag aag aag aaa ggg aag cta cca aga gaa gca aga caa gcg ttg ctc	834
Lys Lys Lys Lys Gly Lys Leu Pro Arg Glu Ala Arg Gln Ala Leu Leu	
235 240 245	
gat tgg tgg aat gtt cat aat aaa tgg cct tac cct act gaa ggc gac	882
Asp Trp Trp Asn Val His Asn Lys Trp Pro Tyr Pro Thr Glu Gly Asp	
250 255 260	
aaa ata gct ctg gct gaa gaa aca ggt ttg gat caa aaa caa atc aac	930
Lys Ile Ala Leu Ala Glu Glu Thr Gly Leu Asp Gln Lys Gln Ile Asn	
265 270 275	
aat tgg ttt ata aac caa agg aaa cgc cat tgg aag cct tcg gag aac	978
Asn Trp Phe Ile Asn Gln Arg Lys Arg His Trp Lys Pro Ser Glu Asn	
280 285 290	
atg ccg ttt gat atg atg gac gat tct aat gaa aca ttc ttt acc gag	1026
Met Pro Phe Asp Met Met Asp Asp Ser Asn Glu Thr Phe Phe Thr Glu	
295 300 305 310	
gaa tga aaagagagac atgggattgt gcattgtata atttttacac tgttttccca	1082
Glu	
agaaaagaaa acagtaaaaa gcttttggtgta aatgggacat catcgcaat gaatggaacc	1142
agttagccaa aacgggtcaag ggcgtggcgt aacgagacat tgtattggaa atagtggcaa	1202
tattatgtca ctaatcttcc aatggtccaa aatgatagat ttcttatttg tattgaacct	1262
tacttagata gctgatgtgt caactaaata atttattttc atccttatac tacttgtatc	1322
aatgtctcta attgatcaat tgttgcttgc tattcaaaaa aaaaaaaaaa aaaaaaa	1379
<210> 8	
<211> 311	
<212> PRT	
<213> Arabidopsis thaliana	
<400> 8	
Met Asp Arg Met Cys Gly Phe Arg Ser Thr Glu Asp Tyr Ser Glu Lys	
1 5 10 15	
Ala Thr Leu Met Met Pro Ser Asp Tyr Gln Ser Leu Ile Cys Ser Thr	

## MBI0018 Sequence Listing.ST25

```

20                               25                               30
Thr Gly Asp Asn Gln Arg Leu Phe Gly Ser Asp Glu Leu Ala Thr Ala
 35                               40                               45

Leu Ser Ser Glu Leu Leu Pro Arg Ile Arg Lys Ala Glu Asp Asn Phe
 50                               55                               60

Ser Leu Ser Val Ile Lys Ser Lys Ile Ala Ser His Pro Leu Tyr Pro
 65                               70                               75                               80

Arg Leu Leu Gln Thr Tyr Ile Asp Cys Gln Lys Val Gly Ala Pro Met
 85                               90                               95

Glu Ile Ala Cys Ile Leu Glu Glu Ile Gln Arg Glu Asn His Val Tyr
100                               105                               110

Lys Arg Asp Val Ala Pro Leu Ser Cys Phe Gly Ala Asp Pro Glu Leu
115                               120                               125

Asp Glu Phe Met Glu Thr Tyr Cys Asp Ile Leu Val Lys Tyr Lys Thr
130                               135                               140

Asp Leu Ala Arg Pro Phe Asp Glu Ala Thr Thr Phe Ile Asn Lys Ile
145                               150                               155                               160

Glu Met Gln Leu Gln Asn Leu Cys Thr Gly Pro Ala Ser Ala Thr Ala
165                               170                               175

Leu Ser Asp Asp Gly Ala Val Ser Ser Asp Glu Glu Leu Arg Glu Asp
180                               185                               190

Asp Asp Ile Ala Ala Asp Asp Ser Gln Gln Arg Ser Asn Asp Arg Asp
195                               200                               205

Leu Lys Asp Gln Leu Leu Arg Lys Phe Gly Ser His Ile Ser Ser Leu
210                               215                               220

Lys Leu Glu Phe Ser Lys Lys Lys Lys Gly Lys Leu Pro Arg Glu
225                               230                               235                               240

Ala Arg Gln Ala Leu Leu Asp Trp Trp Asn Val His Asn Lys Trp Pro
245                               250                               255

Tyr Pro Thr Glu Gly Asp Lys Ile Ala Leu Ala Glu Glu Thr Gly Leu
260                               265                               270

Asp Gln Lys Gln Ile Asn Asn Trp Phe Ile Asn Gln Arg Lys Arg His
275                               280                               285

Trp Lys Pro Ser Glu Asn Met Pro Phe Asp Met Met Asp Asp Ser Asn
290                               295                               300

Glu Thr Phe Phe Thr Glu Glu
305                               310

```

## MBI0018 Sequence Listing.ST25

```

<210> 9
<211> 1571
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (428)..(1402)
<223> G869

<400> 9
aggaacagtg aaaggttcgg ttttttgggg ttcgatctga taatcaacaa gaaaaaaggg 60
tttgatttat gtcggctggg tttgaatcga ctgtgatttt gtctttgatt catatctctt 120
ctccgatttc atcatcatct tcccatcat cgtcgtcttt gaaatcttgt cttctcaacg 180
ctcttcactt ctgctgtaat aagcagagggc ttgttctgga gactccttct ctttccatgc 240
gcttaagacc caaaaggact tgttctagtg ttgaagtctt tggggggttt cacataaagc 300
agcaaaagtt ttcttttttc atagttcgct gagagttttg agttttgata ccaaaaaagt 360
tttgaccttt tagagtgatt ttttgttctt tctgttttct ggggtattttt gaggagtggg 420
ttaaca atg gtt gcg att aga aag gaa cag tct ttg agt ggt gtt agt 469
      Met Val Ala Ile Arg Lys Glu Gln Ser Leu Ser Gly Val Ser
      1 5 10
agc gag att aag aag aga gct aag aga aac act cta tcg tcc ctt cct 517
Ser Glu Ile Lys Lys Arg Ala Lys Arg Asn Thr Leu Ser Ser Leu Pro
15 20 25 30
caa gaa acc caa cct ttg agg aaa gtc cgt att att gtg aat gat cct 565
Gln Glu Thr Gln Pro Leu Arg Lys Val Arg Ile Ile Val Asn Asp Pro
35 40 45
tat gct act gat gat tcc tct agt gat gag gaa gag ctt aag gtt cct 613
Tyr Ala Thr Asp Asp Ser Ser Ser Asp Glu Glu Glu Leu Lys Val Pro
50 55 60
aag cca agg aaa atg aaa cgt atc gtt cgt gag att aac ttt cct tct 661
Lys Pro Arg Lys Met Lys Arg Ile Val Arg Glu Ile Asn Phe Pro Ser
65 70 75
atg gaa gtt tct gaa cag cct tct gag agt tct tct cag gac agt act 709
Met Glu Val Ser Glu Gln Pro Ser Glu Ser Ser Ser Gln Asp Ser Thr
80 85 90
aaa act gat ggc aag ata gct gtg tca gct tct cct gct gtt cct agg 757
Lys Thr Asp Gly Lys Ile Ala Val Ser Ala Ser Pro Ala Val Pro Arg
95 100 105 110
aag aag cct gtt ggt gtt agg caaagg aaa tgg ggg aaa tgg gct gct 805
Lys Lys Pro Val Gly Val Arg Gln Arg Lys Trp Gly Lys Trp Ala Ala
115 120 125
gag att aga gat cct att aag aaa act agg act tgg ttg ggt act ttt 853
Glu Ile Arg Asp Pro Ile Lys Lys Thr Arg Thr Trp Leu Gly Thr Phe
130 135 140
gat act ctt gaa gaa gct gct aaa gct tat gat gct aag aag ctt gag 901
Asp Thr Leu Glu Glu Ala Ala Lys Ala Tyr Asp Ala Lys Lys Leu Glu
145 150 155
ttt gat gct att gtt gct gga aat gtg tcc act act aaa cgt gat gtt 949
Phe Asp Ala Ile Val Ala Gly Asn Val Ser Thr Thr Lys Arg Asp Val
160 165 170
tct tca tct gag act agc caa tgc tct cgt tct tca cct gtt gtt cct 997
Ser Ser Ser Glu Thr Ser Gln Cys Ser Arg Ser Ser Pro Val Val Pro
175 180 185 190
gtt gag caa gat gac act tct gca tca gct ctc act tgt gtc aac aac 1045

```

MBI0018 Sequence Listing.ST25

Val	Glu	Gln	Asp	Asp	Thr	Ser	Ala	Ser	Ala	Leu	Thr	Cys	Val	Asn	Asn		
				195						200				205			
cct	gat	gac	gtc	tcg	acc	gtt	gct	cca	act	gct	cca	act	cca	aat	gtt		1093
Pro	Asp	Asp	Val	Ser	Thr	Val	Ala	Pro	Thr	Ala	Pro	Thr	Pro	Asn	Val		
			210					215					220				
cct	gct	ggg	gga	aac	aag	gaa	acg	ttg	ttc	gat	ttc	gac	ttt	act	aat		1141
Pro	Ala	Gly	Gly	Asn	Lys	Glu	Thr	Leu	Phe	Asp	Phe	Asp	Phe	Thr	Asn		
		225					230					235					
cta	cag	atc	cct	gat	ttt	ggg	ttc	ttg	gca	gag	gag	caa	caa	gac	cta		1189
Leu	Gln	Ile	Pro	Asp	Phe	Gly	Phe	Leu	Ala	Glu	Glu	Gln	Gln	Asp	Leu		
	240					245					250						
gac	ttc	gat	tgt	ttc	ctc	gcg	gat	gat	cag	ttt	gat	gat	ttc	ggc	ttg		1237
Asp	Phe	Asp	Cys	Phe	Leu	Ala	Asp	Asp	Gln	Phe	Asp	Asp	Phe	Gly	Leu		
255					260					265					270		
ctt	gat	gac	att	caa	gga	ttc	gaa	gat	aac	ggg	cca	agt	gcg	tta	cca		1285
Leu	Asp	Asp	Ile	Gln	Gly	Phe	Glu	Asp	Asn	Gly	Pro	Ser	Ala	Leu	Pro		
				275					280					285			
gat	ttc	gac	ttt	gcg	gat	gtt	gaa	gat	ctt	cag	cta	gct	gac	tct	agt		1333
Asp	Phe	Asp	Phe	Ala	Asp	Val	Glu	Asp	Leu	Gln	Leu	Ala	Asp	Ser	Ser		
			290					295					300				
ttc	ggg	ttc	ctt	gat	caa	ctt	gct	cct	atc	aac	atc	tct	tgc	cca	tta		1381
Phe	Gly	Phe	Leu	Asp	Gln	Leu	Ala	Pro	Ile	Asn	Ile	Ser	Cys	Pro	Leu		
	305						310					315					
aaa	agt	ttt	gca	gct	tca	tag	gat	ctt	gctt	agta	atgtta	agtgaga	aaga				1432
Lys	Ser	Phe	Ala	Ala	Ser												
			320														
gtg	ttt	gtt	ttt	cg	ttta	tg	ctt	tag	ta	att	ta	agaca	tac	aaa	agt	gtg	ttccgg
																	1492
att	g	tag	taa	gat	ct	ta	a	ga	ccg	gg	ttt	g	caa	t	tag	ga	atcg
																	1552
aag	ttt	tag	t	at	gt	ttg											1571
<210>	10																
<211>	324																
<212>	PRT																
<213>	Arabidopsis thaliana																
<400>	10																
Met	Val	Ala	Ile	Arg	Lys	Glu	Gln	Ser	Leu	Ser	Gly	Val	Ser	Ser	Glu		
1				5					10					15			
Ile	Lys	Lys	Arg	Ala	Lys	Arg	Asn	Thr	Leu	Ser	Ser	Leu	Pro	Gln	Glu		
			20					25					30				
Thr	Gln	Pro	Leu	Arg	Lys	Val	Arg	Ile	Ile	Val	Asn	Asp	Pro	Tyr	Ala		
		35					40					45					
Thr	Asp	Asp	Ser	Ser	Ser	Asp	Glu	Glu	Glu	Leu	Lys	Val	Pro	Lys	Pro		
	50					55					60						
Arg	Lys	Met	Lys	Arg	Ile	Val	Arg	Glu	Ile	Asn	Phe	Pro	Ser	Met	Glu		
65					70					75					80		
Val	Ser	Glu	Gln	Pro	Ser	Glu	Ser	Ser	Ser	Gln	Asp	Ser	Thr	Lys	Thr		
				85					90					95			
Asp	Gly	Lys	Ile	Ala	Val	Ser	Ala	Ser	Pro	Ala	Val	Pro	Arg	Lys	Lys		
			100					105					110				

## MBI0018 Sequence Listing.ST25

Pro Val Gly Val Arg Gln Arg Lys Trp Gly Lys Trp Ala Ala Glu Ile  
           115                          120                          125  
 Arg Asp Pro Ile Lys Lys Thr Arg Thr Trp Leu Gly Thr Phe Asp Thr  
           130                          135                          140  
 Leu Glu Glu Ala Ala Lys Ala Tyr Asp Ala Lys Lys Leu Glu Phe Asp  
           145                          150                          155                          160  
 Ala Ile Val Ala Gly Asn Val Ser Thr Thr Lys Arg Asp Val Ser Ser  
                           165                          170                          175  
 Ser Glu Thr Ser Gln Cys Ser Arg Ser Ser Pro Val Val Pro Val Glu  
                           180                          185                          190  
 Gln Asp Asp Thr Ser Ala Ser Ala Leu Thr Cys Val Asn Asn Pro Asp  
                           195                          200                          205  
 Asp Val Ser Thr Val Ala Pro Thr Ala Pro Thr Pro Asn Val Pro Ala  
           210                          215                          220  
 Gly Gly Asn Lys Glu Thr Leu Phe Asp Phe Asp Phe Thr Asn Leu Gln  
           225                          230                          235                          240  
 Ile Pro Asp Phe Gly Phe Leu Ala Glu Glu Gln Gln Asp Leu Asp Phe  
                           245                          250                          255  
 Asp Cys Phe Leu Ala Asp Asp Gln Phe Asp Asp Phe Gly Leu Leu Asp  
                           260                          265                          270  
 Asp Ile Gln Gly Phe Glu Asp Asn Gly Pro Ser Ala Leu Pro Asp Phe  
           275                          280                          285  
 Asp Phe Ala Asp Val Glu Asp Leu Gln Leu Ala Asp Ser Ser Phe Gly  
           290                          295                          300  
 Phe Leu Asp Gln Leu Ala Pro Ile Asn Ile Ser Cys Pro Leu Lys Ser  
           305                          310                          315                          320  
 Phe Ala Ala Ser

°  
 <210> 11  
 <211> 1166  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (88)..(951)  
 <223> G1269

<400> 11  
 aacaattctc tctctcttta ttcttcttct tcagcttcag atttcagatc ttaaattcttc 60  
 aagtcttctt cttcttcttc tgcaacc atg gct atg cag gaa cgt tgt gag agt 114  
                           Met Ala Met Gln Glu Arg Cys Glu Ser  
                           1                          5



## MBI0018 Sequence Listing.ST25

tta tgt tct gat gaa ctt ata tct tcc tca gat gcc ttt tac ctc aag	162
Leu Cys Ser Asp Glu Leu Ile Ser Ser Ser Asp Ala Phe Tyr Leu Lys	
10 15 20 25	
aca aga aag cct tat acc atc act aaa caa aga gag aaa tgg aca gaa	210
Thr Arg Lys Pro Tyr Thr Ile Thr Lys Gln Arg Glu Lys Trp Thr Glu	
30 35 40	
gca gag cat gag aag ttt gta gaa gca ttg aaa ctc tat ggc aga gct	258
Ala Glu His Glu Lys Phe Val Glu Ala Leu Lys Leu Tyr Gly Arg Ala	
45 50 55	
tgg aga cga atc gaa gaa cat gtt gga aca aaa act gca gtt cag att	306
Trp Arg Arg Ile Glu Glu His Val Gly Thr Lys Thr Ala Val Gln Ile	
60 65 70	
cga agc cat gcg cag aag ttc ttt act aag gtt gct cgc gat ttt ggt	354
Arg Ser His Ala Gln Lys Phe Phe Thr Lys Val Ala Arg Asp Phe Gly	
75 80 85	
gtt agc tct gag tcc att gag atc ccg cct cca agg cca aag aga aag	402
Val Ser Ser Glu Ser Ile Glu Ile Pro Pro Pro Arg Pro Lys Arg Lys	
90 95 100 105	
ccg atg cat cct tac cct aga aag ctt gtg att cct gat gca aaa gag	450
Pro Met His Pro Tyr Pro Arg Lys Leu Val Ile Pro Asp Ala Lys Glu	
110 115 120	
atg gta tac gct gaa cta acc gga tcc aag ctg att cag gat gaa gat	498
Met Val Tyr Ala Glu Leu Thr Gly Ser Lys Leu Ile Gln Asp Glu Asp	
125 130 135	
aac cga tct cca aca tcg gtt tta tca gct cat ggc tca gat gga tta	546
Asn Arg Ser Pro Thr Ser Val Leu Ser Ala His Gly Ser Asp Gly Leu	
140 145 150	
ggt tcc att ggt tca aat tca cct aac tct tct tca gct gag tta tca	594
Gly Ser Ile Gly Ser Asn Ser Pro Asn Ser Ser Ser Ala Glu Leu Ser	
155 160 165	
tct cac aca gag gaa tca ttg tct cta gaa gca gag acc aaa cag agc	642
Ser His Thr Glu Glu Ser Leu Ser Leu Glu Ala Glu Thr Lys Gln Ser	
170 175 180 185	
ctt aag ctc ttt gga aaa act ttt gta gtt ggt gat tac aac tct tca	690
Leu Lys Leu Phe Gly Lys Thr Phe Val Val Gly Asp Tyr Asn Ser Ser	
190 195 200	
atg agt tgt gat gat tct gaa gat ggc aag aag aag cta tac tca gaa	738
Met Ser Cys Asp Asp Ser Glu Asp Gly Lys Lys Lys Leu Tyr Ser Glu	
205 210 215	
aca cag tct ctt caa tgt tct tct tct act tca gaa aac gct gaa aca	786
Thr Gln Ser Leu Gln Cys Ser Ser Ser Thr Ser Glu Asn Ala Glu Thr	
220 225 230	
gaa gtg gta gtg tcg gag ttc aaa aga agt gag aga tca gct ttc tct	834
Glu Val Val Val Ser Glu Phe Lys Arg Ser Glu Arg Ser Ala Phe Ser	
235 240 245	
cag tta aaa tcg tcg gtg act gag atg aac aac atg aga ggg ttc atg	882
Gln Leu Lys Ser Ser Val Thr Glu Met Asn Asn Met Arg Gly Phe Met	
250 255 260 265	
cct tac aaa aag aga gta aag gtg gaa gaa aac att gac aat gta aaa	930
Pro Tyr Lys Lys Arg Val Lys Val Glu Glu Asn Ile Asp Asn Val Lys	
270 275 280	
tta tca tat cct ttg tgg tga agtgttcgtt tgtgtcaagt cagttgtgta	981
Leu Ser Tyr Pro Leu Trp	
285	
aactctttttg atctcaacat cagattatgt gtataatgtc agagtattag ggaaagtttt	1041

## MBI0018 Sequence Listing.ST25

tttggattag attcgtaaga tcactccaaa gtttcgtgtc tttccatata accagttaga 1101  
aattgagatc cttgtactta aacattttta tttgatcaat caaatcttct tgatgaaaaa 1161  
aaaaa 1166

<210> 12  
<211> 287  
<212> PRT  
<213> Arabidopsis thaliana  
<400> 12

Met Ala Met Gln Glu Arg Cys Glu Ser Leu Cys Ser Asp Glu Leu Ile  
1 5 10 15

Ser Ser Ser Asp Ala Phe Tyr Leu Lys Thr Arg Lys Pro Tyr Thr Ile  
20 25 30

Thr Lys Gln Arg Glu Lys Trp Thr Glu Ala Glu His Glu Lys Phe Val  
35 40 45

Glu Ala Leu Lys Leu Tyr Gly Arg Ala Trp Arg Arg Ile Glu Glu His  
50 55 60

Val Gly Thr Lys Thr Ala Val Gln Ile Arg Ser His Ala Gln Lys Phe  
65 70 75 80

Phe Thr Lys Val Ala Arg Asp Phe Gly Val Ser Ser Glu Ser Ile Glu  
85 90 95

Ile Pro Pro Pro Arg Pro Lys Arg Lys Pro Met His Pro Tyr Pro Arg  
100 105 110

Lys Leu Val Ile Pro Asp Ala Lys Glu Met Val Tyr Ala Glu Leu Thr  
115 120 125

Gly Ser Lys Leu Ile Gln Asp Glu Asp Asn Arg Ser Pro Thr Ser Val  
130 135 140

Leu Ser Ala His Gly Ser Asp Gly Leu Gly Ser Ile Gly Ser Asn Ser  
145 150 155 160

Pro Asn Ser Ser Ser Ala Glu Leu Ser Ser His Thr Glu Glu Ser Leu  
165 170 175

Ser Leu Glu Ala Glu Thr Lys Gln Ser Leu Lys Leu Phe Gly Lys Thr  
180 185 190

Phe Val Val Gly Asp Tyr Asn Ser Ser Met Ser Cys Asp Asp Ser Glu  
195 200 205

Asp Gly Lys Lys Lys Leu Tyr Ser Glu Thr Gln Ser Leu Gln Cys Ser  
210 215 220

Ser Ser Thr Ser Glu Asn Ala Glu Thr Glu Val Val Val Ser Glu Phe  
225 230 235 240

Lys Arg Ser Glu Arg Ser Ala Phe Ser Gln Leu Lys Ser Ser Val Thr

## MBI0018 Sequence Listing.ST25

245 250 255

Glu Met Asn Asn Met Arg Gly Phe Met Pro Tyr Lys Lys Arg Val Lys  
260 265 270

Val Glu Glu Asn Ile Asp Asn Val Lys Leu Ser Tyr Pro Leu Trp  
275 280 285

<210> 13  
<211> 2031  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (240)..(1574)  
<223> G1038

<400> 13  
gctcgttttc aaattaaaaa cagggagaaa tttggaaatt ccagtacgac gggagataaa 60  
acctaacata cgccatgggtg accgttatct aaactacgcc aaaatatttg aagtgtcgtc 120  
gtttcataat aaaacgcaaa caaaaaccca ctcccacttt ctcctttcca aaaaaagaac 180  
tctcgccact ttctctgctc ttttctttct ctctctcttt cttgttttcg ccggcgatc 239  
atg gag aaa agc ggc ttc tct ccc gtc ggt cta agg gtt ctt gtc gta 287  
Met Glu Lys Ser Gly Phe Ser Pro Val Gly Leu Arg Val Leu Val Val  
1 5 10 15  
gac gat gat cca act tgg ctc aag att ctc gag aaa atg ctc aag aag 335  
Asp Asp Asp Pro Thr Trp Leu Lys Ile Leu Glu Lys Met Leu Lys Lys  
20 25 30  
tgt tct tac gaa gta acg acc tgt gga tta gct aga gag gct ttg agg 383  
Cys Ser Tyr Glu Val Thr Thr Cys Gly Leu Ala Arg Glu Ala Leu Arg  
35 40 45  
ttg ctg agg gag cgt aaa gat gga tat gat atc gtg atc agc gat gtg 431  
Leu Leu Arg Glu Arg Lys Asp Gly Tyr Asp Ile Val Ile Ser Asp Val  
50 55 60  
aac atg cct gac atg gat ggt ttc aag ctt ctt gag cat gtt ggt ctt 479  
Asn Met Pro Asp Met Asp Gly Phe Lys Leu Leu Glu His Val Gly Leu  
65 70 75 80  
gaa tta gac ctc cct gta ata atg atg tgc gtg gac ggc gaa aca agc 527  
Glu Leu Asp Leu Pro Val Ile Met Met Ser Val Asp Gly Glu Thr Ser  
85 90 95  
cga gtg atg aag gga gtg cac acg gga gct tgt gat tac ctc ttg aag 575  
Arg Val Met Lys Gly Val His Thr Gly Ala Cys Asp Tyr Leu Leu Lys  
100 105 110  
ccg ata aga atg aag gag tta aag att ata tgg caa cat gtt ctg aga 623  
Pro Ile Arg Met Lys Glu Leu Lys Ile Ile Trp Gln His Val Leu Arg  
115 120 125  
aag aag ctt caa gaa gtg aga gat atc gaa ggc tgt gga tac gaa gga 671  
Lys Lys Leu Gln Glu Val Arg Asp Ile Glu Gly Cys Gly Tyr Glu Gly  
130 135 140  
gga gcg gat tgg atc act cga tac gat gaa gca cat ttt ctt gga ggt 719  
Gly Ala Asp Trp Ile Thr Arg Tyr Asp Glu Ala His Phe Leu Gly Gly  
145 150 155 160  
ggt gaa gat gtt tct ttt ggg aaa aag aga aaa gac ttt gac ttt gag 767  
Gly Glu Asp Val Ser Phe Gly Lys Lys Arg Lys Asp Phe Asp Phe Glu  
165 170 175

MBI0018 Sequence Listing.ST25																
aag aag ctt ctt caa gat gag agt gat cca tca tct tct tct tcc aag	Lys Lys Leu Leu Gln Asp Glu Ser Asp Pro Ser Ser Ser Ser Ser Lys	180	185	190												815
aaa gct aga gtt gtt tgg tct ttt gag ctt cat cat aag ttt gtc aac	Lys Ala Arg Val Val Trp Ser Phe Glu Leu His His Lys Phe Val Asn	195	200	205												863
gcc gtt aac caa atc gga tgc gat cac aaa gct ggt ccc aag aag ata	Ala Val Asn Gln Ile Gly Cys Asp His Lys Ala Gly Pro Lys Lys Ile	210	215	220												911
ttg gat ctc atg aat gtt cca tgg ctc act aga gaa aat gtt gca agc	Leu Asp Leu Met Asn Val Pro Trp Leu Thr Arg Glu Asn Val Ala Ser	225	230	235												959
cac ctt cag aaa tat aga ctt tac ctg agc aga tta gag aaa gga aag	His Leu Gln Lys Tyr Arg Leu Tyr Leu Ser Arg Leu Glu Lys Gly Lys	245	250	255												1007
gag ctc aag tgt tat tca ggt ggc gtg aag aat gcg gat tca tct cca	Glu Leu Lys Cys Tyr Ser Gly Gly Val Lys Asn Ala Asp Ser Ser Pro	260	265	270												1055
aaa gat gtc gaa gtg aat tca ggc tac caa agc cct ggg agg agc agc	Lys Asp Val Glu Val Asn Ser Gly Tyr Gln Ser Pro Gly Arg Ser Ser	275	280	285												1103
tat gta ttc tct gga gga aat tct ctg atc caa aaa gca aca gag att	Tyr Val Phe Ser Gly Gly Asn Ser Leu Ile Gln Lys Ala Thr Glu Ile	290	295	300												1151
gat cca aag cca ctt gct tca gct tct ttg tct gac ccc aac acc gat	Asp Pro Lys Pro Leu Ala Ser Ala Ser Leu Ser Asp Pro Asn Thr Asp	305	310	315												1199
gtg atc atg cct ccg aaa aca aaa aag acg cgt ata gga ttt gat cct	Val Ile Met Pro Lys Thr Lys Lys Thr Arg Ile Gly Phe Asp Pro	325	330	335												1247
ccc att tcc tcc tct gcg ttt gac tct ctg ctt cct tgg aat gat gtt	Pro Ile Ser Ser Ser Ala Phe Asp Ser Leu Leu Pro Trp Asn Asp Val	340	345	350												1295
cca gag gtc ctt gaa tcg aag ccg gtt ctg tat gag aat agc ttt ctc	Pro Glu Val Leu Glu Ser Lys Pro Val Leu Tyr Glu Asn Ser Phe Leu	355	360	365												1343
cag caa caa cca ttg cca agt caa agt tcc tat gtt gca att tct gca	Gln Gln Gln Pro Leu Pro Ser Gln Ser Ser Tyr Val Ala Ile Ser Ala	370	375	380												1391
cca tct ctc atg gag gag gaa atg aag cct cct tat gag aca cca gca	Pro Ser Leu Met Glu Glu Glu Met Lys Pro Pro Tyr Glu Thr Pro Ala	385	390	395												1439
gga ggc agt agt gtg aat gca gat gag ttt ctc atg cca caa gac aag	Gly Gly Ser Ser Val Asn Ala Asp Glu Phe Leu Met Pro Gln Asp Lys	405	410	415												1487
atc cct act gta acc ctt caa gat ttg gat ccc tct gcc atg aag ctg	Ile Pro Thr Val Thr Leu Gln Asp Leu Asp Pro Ser Ala Met Lys Leu	420	425	430												1535
cag gag ttc aac aca gaa ggc gat tct gaa gaa gct tga actggggaac	Gln Glu Phe Asn Thr Glu Gly Asp Ser Glu Glu Ala	435	440													1584
ttccagaatc acatcattct gtttcttttag acactgactt agacttgact tggcttcaag																1644
gcgagcggtt cttgcaaaca ccgactccag tttcaagata cagtagtagc ccatcactcc																1704
tatctgagct cccagcccac cttaattggt atggaaatga gcggctgcct gaccctgacg																1764
agtattcctt catggtagac caaggtttat tcatatctta accttggtcc aataacttct																1824

## MBI0018 Sequence Listing.ST25

```

tttcgtatat tggttggtgt aatgcagaaa gattttgtgg gtatacctga aaataatctt 1884
gctttcccaa gaaccttcca tgatcggatg cattgtacaa taatccacga gtgtcgtagg 1944
ctaattacac caaacagggt gatgacagtg ataaggccac atgtttcaca ccgtcgctta 2004
agatctttac tgtcacctgg aaggaaa 2031

```

```

<210> 14
<211> 444
<212> PRT
<213> Arabidopsis thaliana

```

```

<400> 14

```

```

Met Glu Lys Ser Gly Phe Ser Pro Val Gly Leu Arg Val Leu Val Val
1      5      10      15

Asp Asp Asp Pro Thr Trp Leu Lys Ile Leu Glu Lys Met Leu Lys Lys
20      25      30

Cys Ser Tyr Glu Val Thr Thr Cys Gly Leu Ala Arg Glu Ala Leu Arg
35      40      45

Leu Leu Arg Glu Arg Lys Asp Gly Tyr Asp Ile Val Ile Ser Asp Val
50      55      60

Asn Met Pro Asp Met Asp Gly Phe Lys Leu Leu Glu His Val Gly Leu
65      70      75      80

Glu Leu Asp Leu Pro Val Ile Met Met Ser Val Asp Gly Glu Thr Ser
85      90      95

Arg Val Met Lys Gly Val His Thr Gly Ala Cys Asp Tyr Leu Leu Lys
100     105     110

Pro Ile Arg Met Lys Glu Leu Lys Ile Ile Trp Gln His Val Leu Arg
115     120     125

Lys Lys Leu Gln Glu Val Arg Asp Ile Glu Gly Cys Gly Tyr Glu Gly
130     135     140

Gly Ala Asp Trp Ile Thr Arg Tyr Asp Glu Ala His Phe Leu Gly Gly
145     150     155     160

Gly Glu Asp Val Ser Phe Gly Lys Lys Arg Lys Asp Phe Asp Phe Glu
165     170     175

Lys Lys Leu Leu Gln Asp Glu Ser Asp Pro Ser Ser Ser Ser Ser Lys
180     185     190

Lys Ala Arg Val Val Trp Ser Phe Glu Leu His His Lys Phe Val Asn
195     200     205

Ala Val Asn Gln Ile Gly Cys Asp His Lys Ala Gly Pro Lys Lys Ile
210     215     220

Leu Asp Leu Met Asn Val Pro Trp Leu Thr Arg Glu Asn Val Ala Ser
225     230     235     240

```

MBI0018 Sequence Listing.ST25

His	Leu	Gln	Lys	Tyr 245	Arg	Leu	Tyr	Leu	Ser 250	Arg	Leu	Glu	Lys	Gly 255	Lys
Glu	Leu	Lys	Cys 260	Tyr	Ser	Gly	Gly	Val 265	Lys	Asn	Ala	Asp	Ser 270	Ser	Pro
Lys	Asp	Val 275	Glu	Val	Asn	Ser	Gly 280	Tyr	Gln	Ser	Pro	Gly 285	Arg	Ser	Ser
Tyr	Val 290	Phe	Ser	Gly	Gly	Asn 295	Ser	Leu	Ile	Gln	Lys 300	Ala	Thr	Glu	Ile
Asp 305	Pro	Lys	Pro	Leu	Ala 310	Ser	Ala	Ser	Leu	Ser 315	Asp	Pro	Asn	Thr	Asp 320
Val	Ile	Met	Pro	Pro 325	Lys	Thr	Lys	Lys	Thr 330	Arg	Ile	Gly	Phe	Asp 335	Pro
Pro	Ile	Ser	Ser 340	Ser	Ala	Phe	Asp	Ser 345	Leu	Leu	Pro	Trp	Asn 350	Asp	Val
Pro	Glu	Val 355	Leu	Glu	Ser	Lys	Pro 360	Val	Leu	Tyr	Glu	Asn 365	Ser	Phe	Leu
Gln 370	Gln	Gln	Pro	Leu	Pro	Ser 375	Gln	Ser	Ser	Tyr	Val 380	Ala	Ile	Ser	Ala
Pro 385	Ser	Leu	Met	Glu	Glu 390	Glu	Met	Lys	Pro	Pro 395	Tyr	Glu	Thr	Pro	Ala 400
Gly	Gly	Ser	Ser	Val 405	Asn	Ala	Asp	Glu	Phe 410	Leu	Met	Pro	Gln	Asp 415	Lys
Ile	Pro	Thr	Val 420	Thr	Leu	Gln	Asp	Leu 425	Asp	Pro	Ser	Ala	Met 430	Lys	Leu
Gln	Glu	Phe 435	Asn	Thr	Glu	Gly	Asp 440	Ser	Glu	Glu	Ala				

```
<210> 15
<211> 2821
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>
<221> CDS
<222> (188) .. (2716)
<223> G438
```

```
<400> 15
cggggtaccc aagccacgac cgtagaatct tcttttgtct gaaaagaatt acaattttacg 60
tttctcttac gatacgacgg actttccgaa gaaattaatt taaagagaaa agaagaagaa 120
gccaaagaag aagaagaagc tagaagaaac agtaaagttt gagacttttt ttgagggtcg 180
agctaaa atg gag atg gcg gtg gct aac cac cgt gag aga agc agt gac 229
Met Glu Met Ala Val Ala Asn His Arg Glu Arg Ser Ser Asp
1 5 10
```

## MBI0018 Sequence Listing.ST25

agt atg aat aga cat tta gat agt agc ggt aag tac gtt agg tac aca	277
Ser Met Asn Arg His Leu Asp Ser Ser Gly Lys Tyr Val Arg Tyr Thr	
15 20 25 30	
gct gag caa gtc gag gct ctt gag cgt gtc tac gct gag tgt cct aag	325
Ala Glu Gln Val Glu Ala Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys	
35 40 45	
cct agc tct ctc cgt cga caa caa ttg atc cgt gaa tgt tcc att ttg	373
Pro Ser Ser Leu Arg Arg Gln Gln Ile Arg Glu Cys Ser Ile Leu	
50 55 60	
gcc aat att gag cct aag cag atc aaa gtc tgg ttt cag aac cgc agg	421
Ala Asn Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg	
65 70 75	
tgt cga gat aag cag agg aaa gag gcg tcg agg ctc cag agc gta aac	469
Cys Arg Asp Lys Gln Arg Lys Glu Ala Ser Arg Leu Gln Ser Val Asn	
80 85 90	
cgg aag ctc tct gcg atg aat aaa ctg ttg atg gag gag aat gat agg	517
Arg Lys Leu Ser Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg	
95 100 105 110	
ttg cag aag cag gtt tct cag ctt gtc tgc gaa aat gga tat atg aaa	565
Leu Gln Lys Gln Val Ser Gln Leu Val Cys Glu Asn Gly Tyr Met Lys	
115 120 125	
cag cag cta act act gtt gtt aac gat cca agc tgt gaa tct gtg gtc	613
Gln Gln Leu Thr Thr Val Val Asn Asp Pro Ser Cys Glu Ser Val Val	
130 135 140	
aca act cct cag cat tcg ctt aga gat gcg aat agt cct gct gga ttg	661
Thr Thr Pro Gln His Ser Leu Arg Asp Ala Asn Ser Pro Ala Gly Leu	
145 150 155	
ctc tca atc gca gag gag act ttg gca gag ttc cta tcc aag gct aca	709
Leu Ser Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala Thr	
160 165 170	
gga act gct gtt gat tgg gtt cag atg cct ggg atg aag cct ggt ccg	757
Gly Thr Ala Val Asp Trp Val Gln Met Pro Gly Met Lys Pro Gly Pro	
175 180 185 190	
gat tcg gtt ggc atc ttt gcc att tcg caa aga tgc aat gga gtg gca	805
Asp Ser Val Gly Ile Phe Ala Ile Ser Gln Arg Cys Asn Gly Val Ala	
195 200 205	
gct cga gcc tgt ggt ctt gtt agc tta gaa cct atg aag att gca gag	853
Ala Arg Ala Cys Gly Leu Val Ser Leu Glu Pro Met Lys Ile Ala Glu	
210 215 220	
atc ctc aaa gat cgg cca tct tgg ttc cgt gac tgt agg agc ctt gaa	901
Ile Leu Lys Asp Arg Pro Ser Trp Phe Arg Asp Cys Arg Ser Leu Glu	
225 230 235	
gtt ttc act atg ttc ccg gct ggt aat ggt ggc aca atc gag ctt gtt	949
Val Phe Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu Leu Val	
240 245 250	
tat atg cag acg tat gca cca acg act ctg gct cct gcc cgc gat ttc	997
Tyr Met Gln Thr Tyr Ala Pro Thr Thr Leu Ala Pro Ala Arg Asp Phe	
255 260 265 270	
tgg acc ctg aga tac aca acg agc ctc gac aat ggg agt ttt gtg gtt	1045
Trp Thr Leu Arg Tyr Thr Thr Ser Leu Asp Asn Gly Ser Phe Val Val	
275 280 285	
tgt gag agg tcg cta tct ggc tct gga gct ggg cct aat gct gct tca	1093
Cys Glu Arg Ser Leu Ser Gly Ser Gly Ala Gly Pro Asn Ala Ala Ser	
290 295 300	
gct tct cag ttt gtg aga gca gaa atg ctt tct agt ggg tat tta ata	1141
Ala Ser Gln Phe Val Arg Ala Glu Met Leu Ser Ser Gly Tyr Leu Ile	

MBI0018 Sequence Listing.ST25																	
305				310				315									
agg Arg	cct Pro 320	tgt Cys	gat Asp	ggt Gly	ggt Gly	ggt Gly 325	tct Ser	att Ile	att Ile	cac His	att Ile 330	gtc Val	gat Asp	cac His	ctt Leu		1189
aat Asn 335	ctt Leu	gag Glu	gct Ala	tgg Trp	agt Ser 340	gtt Val	ccg Pro	gat Asp	gtg Val	ctt Leu 345	cga Arg	ccc Pro	ctt Leu	tat Tyr	gag Glu 350		1237
tca Ser	tcc Ser	aaa Lys	gtc Val 355	gtt Val	gca Ala	caa Gln	aaa Lys	atg Met	acc Thr 360	att Ile	tcc Ser	gcg Ala	ttg Leu	cgg Arg 365	tat Tyr		1285
atc Ile	agg Arg	caa Gln	tta Leu 370	gcc Ala	caa Gln	gag Glu	tct Ser	aat Asn 375	ggt Gly	gaa Glu	gta Val	gtg Val	tat Tyr 380	gga Gly	tta Leu		1333
gga Gly	agg Arg	cag Gln 385	cct Pro	gct Ala	gtt Val	ctt Leu	aga Arg 390	acc Thr	ttt Phe	agc Ser	caa Gln 395	aga Arg	tta Leu	agc Ser	agg Arg		1381
ggc Gly	ttc Phe 400	aat Asn	gat Asp	gcg Ala	gtt Val	aat Asn 405	ggg Gly	ttt Phe	ggt Gly	gac Asp 410	gac Asp	ggg Gly	tgg Trp	tct Ser	acg Thr		1429
atg Met 415	cat His	tgt Cys	gat Asp	gga Gly	gcg Ala 420	gaa Glu	gat Asp	att Ile	atc Ile	gtt Val 425	gct Ala	att Ile	aac Asn	tct Ser	aca Thr 430		1477
aag Lys	cat His	ttg Leu	aat Asn	aat Asn 435	att Ile	tct Ser	aat Asn	tct Ser	ctt Leu 440	tcg Ser	ttc Phe	ctt Leu	gga Gly	ggc Gly 445	gtg Val		1525
ctc Leu	tgt Cys	gcc Ala	aag Lys 450	gct Ala	tca Ser	atg Met	ctt Leu	ctc Leu 455	caa Gln	aat Asn	gtt Val	cct Pro	cct Pro 460	gcg Ala	gtt Val		1573
ttg Leu	atc Ile	cgg Arg 465	ttc Phe	ctt Leu	aga Arg	gag Glu	cat His 470	cga Arg	tct Ser	gag Glu	tgg Trp	gct Ala 475	gat Asp	ttc Phe	aat Asn		1621
gtt Val	gat Asp 480	gca Ala	tat Tyr	tcc Ser	gct Ala	gct Ala 485	aca Thr	ctt Leu	aaa Lys	gct Ala 490	ggt Gly	agc Ser	ttt Phe	gct Ala	tat Tyr		1669
ccg Pro 495	gga Gly	atg Met	aga Arg	cca Pro	aca Thr 500	aga Arg	ttc Phe	act Thr	ggg Gly	agt Ser 505	cag Gln	atc Ile	ata Ile	atg Met	cca Pro 510		1717
cta Leu	gga Gly	cat His	aca Thr	att Ile 515	gaa Glu	cac His	gaa Glu	gaa Glu 520	atg Met	cta Leu	gaa Glu	gtt Val	gtt Val	aga Arg 525	ctg Leu		1765
gaa Glu	ggt Gly	cat His	tct Ser 530	ctt Leu	gct Ala	caa Gln	gaa Glu	gat Asp 535	gca Ala	ttt Phe	atg Met	tca Ser	cgg Arg 540	gat Asp	gtc Val		1813
cat His	ctc Leu	ctt Leu 545	cag Gln	att Ile	tgt Cys	acc Thr	ggg Gly 550	att Ile	gac Asp	gag Glu	aat Asn	gcc Ala 555	gtt Val	gga Gly	gct Ala		1861
tgt Cys	tct Ser 560	gaa Glu	ctg Leu	ata Ile	ttt Phe	gct Ala 565	ccg Pro	att Ile	aat Asn	gag Glu	atg Met 570	ttc Phe	ccg Pro	gat Asp	gat Asp		1909
gct Ala 575	cca Pro	ctt Leu	gtt Val	ccc Pro	tct Ser 580	gga Gly	ttc Phe	cga Arg	gtc Val	ata Ile 585	ccc Pro	gtt Val	gat Asp	gct Ala	aaa Lys 590		1957
acg Thr	gga Gly	gat Asp	gta Val	caa Gln 595	gat Asp	ctg Leu	tta Leu	acc Thr	gct Ala 600	aat Asn	cac His	cgt Arg	aca Thr	cta Leu 605	gac Asp		2005
tta	act	tct	agc	ctt	gaa	gtc	ggt	cca	tca	cct	gag	aat	gct	tct	gga		2053



MBI0018 Sequence Listing.ST25

Leu Thr Ser Ser 610	Leu Glu Val Gly Pro 615	Ser Pro Glu Asn Ala Ser Gly 620	
aac tct ttt tct agc tca agc tcg aga tgt att ctc act atc gcg ttt	Asn Ser Phe Ser Ser Ser Ser Ser Arg Cys Ile Leu Thr Ile Ala Phe		2101
caa ttc cct ttt gaa aac aac ttg caa gaa aat gtt gct ggt atg gct	Gln Phe Pro Phe Glu Asn Asn Leu Gln Glu Asn Val Ala Gly Met Ala		2149
tgt cag tat gtg agg agc gtg atc tca tca gtt caa cgt gtt gca atg	Cys Gln Tyr Val Arg Ser Val Ile Ser Ser Val Gln Arg Val Ala Met		2197
gcg atc tca ccg tct ggg ata agc ccg agt ctg ggc tcc aaa ttg tcc	Ala Ile Ser Pro Ser Gly Ile Ser Pro Ser Leu Gly Ser Lys Leu Ser		2245
cca gga tct cct gaa gct gtt act ctt gct cag tgg atc tct caa agt	Pro Gly Ser Pro Glu Ala Val Thr Leu Ala Gln Trp Ile Ser Gln Ser		2293
tac agt cat cac tta ggc tcg gag ttg ctg acg att gat tca ctt gga	Tyr Ser His His Leu Gly Ser Glu Leu Leu Thr Ile Asp Ser Leu Gly		2341
agc gac gac tcg gta cta aaa ctt cta tgg gat cac caa gat gcc atc	Ser Asp Asp Ser Val Leu Lys Leu Leu Trp Asp His Gln Asp Ala Ile		2389
ctg tgt tgc tca tta aag cca cag cca gtg ttc atg ttt gcg aac caa	Leu Cys Cys Ser Leu Lys Pro Gln Pro Val Phe Met Phe Ala Asn Gln		2437
gct ggt cta gac atg cta gag aca aca ctt gta gcc tta caa gat ata	Ala Gly Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp Ile		2485
aca ctc gaa aag ata ttc gat gaa tcg ggt cgt aag gct atc tgt tcg	Thr Leu Glu Lys Ile Phe Asp Glu Ser Gly Arg Lys Ala Ile Cys Ser		2533
gac ttc gcc aag cta atg caa cag gga ttt gct tgc ttg cct tca gga	Asp Phe Ala Lys Leu Met Gln Gln Gly Phe Ala Cys Leu Pro Ser Gly		2581
atc tgt gtg tca acg atg gga aga cat gtg agt tat gaa caa gct gtt	Ile Cys Val Ser Thr Met Gly Arg His Val Ser Tyr Glu Gln Ala Val		2629
gct tgg aaa gtg ttt gct gca tct gaa gaa aac aac aac aat ctg cat	Ala Trp Lys Val Phe Ala Ala Ser Glu Glu Asn Asn Asn Asn Leu His		2677
tgt ctt gcc ttc tcc ttt gta aac tgg tct ttt gtg tga ttcgattgac	Cys Leu Ala Phe Ser Phe Val Asn Trp Ser Phe Val		2726
agaaaaagac taatttaaatt tacgtaga gaactcaaat ttttggtgtg tgttaggtg			2786
tctctgtttt gttttttaaa attattttga tcaaa			2821
<210> 16			
<211> 842			
<212> PRT			
<213> Arabidopsis thaliana			
<400> 16			
Met Glu Met Ala Val Ala Asn His Arg Glu Arg Ser Ser Asp Ser Met			
1	5	10	15

## MBI0018 Sequence Listing.ST25

```

Asn Arg His Leu Asp Ser Ser Gly Lys Tyr Val Arg Tyr Thr Ala Glu
20          25          30

Gln Val Glu Ala Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys Pro Ser
35          40          45

Ser Leu Arg Arg Gln Gln Leu Ile Arg Glu Cys Ser Ile Leu Ala Asn
50          55          60

Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg
65          70          75          80

Asp Lys Gln Arg Lys Glu Ala Ser Arg Leu Gln Ser Val Asn Arg Lys
85          90          95

Leu Ser Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln
100         105         110

Lys Gln Val Ser Gln Leu Val Cys Glu Asn Gly Tyr Met Lys Gln Gln
115         120         125

Leu Thr Thr Val Val Asn Asp Pro Ser Cys Glu Ser Val Val Thr Thr
130         135         140

Pro Gln His Ser Leu Arg Asp Ala Asn Ser Pro Ala Gly Leu Leu Ser
145         150         155         160

Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala Thr Gly Thr
165         170         175

Ala Val Asp Trp Val Gln Met Pro Gly Met Lys Pro Gly Pro Asp Ser
180         185         190

Val Gly Ile Phe Ala Ile Ser Gln Arg Cys Asn Gly Val Ala Ala Arg
195         200         205

Ala Cys Gly Leu Val Ser Leu Glu Pro Met Lys Ile Ala Glu Ile Leu
210         215         220

Lys Asp Arg Pro Ser Trp Phe Arg Asp Cys Arg Ser Leu Glu Val Phe
225         230         235         240

Thr Met Phe Pro Ala Gly Asn Gly Gly Thr Ile Glu Leu Val Tyr Met
245         250         255

Gln Thr Tyr Ala Pro Thr Thr Leu Ala Pro Ala Arg Asp Phe Trp Thr
260         265         270

Leu Arg Tyr Thr Thr Ser Leu Asp Asn Gly Ser Phe Val Val Cys Glu
275         280         285

Arg Ser Leu Ser Gly Ser Gly Ala Gly Pro Asn Ala Ala Ser Ala Ser
290         295         300

Gln Phe Val Arg Ala Glu Met Leu Ser Ser Gly Tyr Leu Ile Arg Pro
305         310         315         320

```

## MBI0018 Sequence Listing.ST25

Cys Asp Gly Gly Gly Ser Ile Ile His Ile Val Asp His Leu Asn Leu  
 325 330 335  
 Glu Ala Trp Ser Val Pro Asp Val Leu Arg Pro Leu Tyr Glu Ser Ser  
 340 345 350  
 Lys Val Val Ala Gln Lys Met Thr Ile Ser Ala Leu Arg Tyr Ile Arg  
 355 360 365  
 Gln Leu Ala Gln Glu Ser Asn Gly Glu Val Val Tyr Gly Leu Gly Arg  
 370 375 380  
 Gln Pro Ala Val Leu Arg Thr Phe Ser Gln Arg Leu Ser Arg Gly Phe  
 385 390 395 400  
 Asn Asp Ala Val Asn Gly Phe Gly Asp Asp Gly Trp Ser Thr Met His  
 405 410 415  
 Cys Asp Gly Ala Glu Asp Ile Ile Val Ala Ile Asn Ser Thr Lys His  
 420 425 430  
 Leu Asn Asn Ile Ser Asn Ser Leu Ser Phe Leu Gly Gly Val Leu Cys  
 435 440 445  
 Ala Lys Ala Ser Met Leu Leu Gln Asn Val Pro Pro Ala Val Leu Ile  
 450 455 460  
 Arg Phe Leu Arg Glu His Arg Ser Glu Trp Ala Asp Phe Asn Val Asp  
 465 470 475 480  
 Ala Tyr Ser Ala Ala Thr Leu Lys Ala Gly Ser Phe Ala Tyr Pro Gly  
 485 490 495  
 Met Arg Pro Thr Arg Phe Thr Gly Ser Gln Ile Ile Met Pro Leu Gly  
 500 505 510  
 His Thr Ile Glu His Glu Glu Met Leu Glu Val Val Arg Leu Glu Gly  
 515 520 525  
 His Ser Leu Ala Gln Glu Asp Ala Phe Met Ser Arg Asp Val His Leu  
 530 535 540  
 Leu Gln Ile Cys Thr Gly Ile Asp Glu Asn Ala Val Gly Ala Cys Ser  
 545 550 555 560  
 Glu Leu Ile Phe Ala Pro Ile Asn Glu Met Phe Pro Asp Asp Ala Pro  
 565 570 575  
 Leu Val Pro Ser Gly Phe Arg Val Ile Pro Val Asp Ala Lys Thr Gly  
 580 585 590  
 Asp Val Gln Asp Leu Leu Thr Ala Asn His Arg Thr Leu Asp Leu Thr  
 595 600 605  
 Ser Ser Leu Glu Val Gly Pro Ser Pro Glu Asn Ala Ser Gly Asn Ser  
 610 615 620

## MBI0018 Sequence Listing.ST25

Phe Ser Ser Ser Ser Ser Arg Cys Ile Leu Thr Ile Ala Phe Gln Phe  
 625 630 635 640  
 Pro Phe Glu Asn Asn Leu Gln Glu Asn Val Ala Gly Met Ala Cys Gln  
 645 650 655  
 Tyr Val Arg Ser Val Ile Ser Ser Val Gln Arg Val Ala Met Ala Ile  
 660 665 670  
 Ser Pro Ser Gly Ile Ser Pro Ser Leu Gly Ser Lys Leu Ser Pro Gly  
 675 680 685  
 Ser Pro Glu Ala Val Thr Leu Ala Gln Trp Ile Ser Gln Ser Tyr Ser  
 690 695 700  
 His His Leu Gly Ser Glu Leu Leu Thr Ile Asp Ser Leu Gly Ser Asp  
 705 710 715 720  
 Asp Ser Val Leu Lys Leu Leu Trp Asp His Gln Asp Ala Ile Leu Cys  
 725 730 735  
 Cys Ser Leu Lys Pro Gln Pro Val Phe Met Phe Ala Asn Gln Ala Gly  
 740 745 750  
 Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp Ile Thr Leu  
 755 760 765  
 Glu Lys Ile Phe Asp Glu Ser Gly Arg Lys Ala Ile Cys Ser Asp Phe  
 770 775 780  
 Ala Lys Leu Met Gln Gln Gly Phe Ala Cys Leu Pro Ser Gly Ile Cys  
 785 790 795 800  
 Val Ser Thr Met Gly Arg His Val Ser Tyr Glu Gln Ala Val Ala Trp  
 805 810 815  
 Lys Val Phe Ala Ala Ser Glu Glu Asn Asn Asn Leu His Cys Leu  
 820 825 830  
 Ala Phe Ser Phe Val Asn Trp Ser Phe Val  
 835 840

<210> 17  
 <211> 1888  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (326)..(1708)  
 <223> G571

<400> 17  
 tagccgacct ctcttctctc ttctgaaaaa aacaccaaag gagctttaa tgctccgtta 60  
 cataatctct atctctttcc aagaatatag agaaaggaaa ataatatata agaattaaaa 120  
 gaaggatat catcatctct ctagctagtg atcaaagcac cgatcatcatc atcatatata 180

## MBI0018 Sequence Listing.ST25

atcagcttgc	ctcagaggag	aagaccaaca	taagagagat	cgaagatcaa	aatctatctc	240						
tcttcatcat	cttctgctgt	tactatcata	tcacacgctc	tctcaaacat	catcctatat	300						
atagacttct	cttcatcatc	atcaa	atg	caa	ggg	cat	cac	cag	aat	cat	cat	352
		Met	Gln	Gly	His	His	Gln	Asn	His	His		
		1			5							
caa	cac	tta	tca	tca	tcc	tcc	gcc	acg	tct	tcc	cat	400
Gln	His	Leu	Ser	Ser	Ser	Ser	Ala	Thr	Ser	Ser	His	
10					15				20			
aac	aaa	gat	ggg	tat	gat	att	gga	gag	ata	gac	cca	448
Asn	Lys	Asp	Gly	Tyr	Asp	Ile	Gly	Glu	Ile	Asp	Pro	
				30					35			
tat	ctt	gat	gga	caa	gga	cat	cat	gat	cct	cca	tca	496
Tyr	Leu	Asp	Gly	Gln	Gly	His	His	Asp	Pro	Pro	Ser	
			45					50				
cct	tta	cat	cat	cat	cac	aca	act	cag	aat	ttg	gcg	544
Pro	Leu	His	His	His	His	Thr	Thr	Gln	Asn	Leu	Ala	
aca	tcg	acg	ctc	aac	atc	ttt	cca	tct	cag	cct	atg	592
Thr	Ser	Thr	Leu	Asn	Ile	Phe	Pro	Ser	Gln	Pro	Met	
cct	cct	tct	tct	aca	cac	aat	acc	gat	aat	aca	aga	640
Pro	Pro	Ser	Ser	Thr	His	Asn	Thr	Asp	Asn	Thr	Arg	
gct	caa	cct	agt	ggg	tcc	act	cga	cca	gct	tct	gac	688
Ala	Gln	Pro	Ser	Gly	Ser	Thr	Arg	Pro	Ala	Ser	Asp	
ttg	acc	aat	cat	tct	cag	ttt	cat	caa	cct	cct	caa	736
Leu	Thr	Asn	His	Ser	Gln	Phe	His	Gln	Pro	Pro	Gln	
atc	aag	aag	gaa	ggg	aac	cgc	aag	ggg	ctt	gcc	tca	784
Ile	Lys	Lys	Glu	Gly	Asn	Arg	Lys	Gly	Leu	Ala	Ser	
ata	cct	aaa	tcg	tca	gac	cct	aaa	aca	ttg	aga	aga	832
Ile	Pro	Lys	Ser	Ser	Asp	Pro	Lys	Thr	Leu	Arg	Arg	
aga	gaa	gca	gca	aga	aaa	agc	aga	tta	cgt	aaa	aag	880
Arg	Glu	Ala	Ala	Arg	Lys	Ser	Arg	Leu	Arg	Lys	Lys	
caa	ctc	gag	tca	tgt	agg	atc	aaa	ctg	acc	caa	cta	928
Gln	Leu	Glu	Ser	Cys	Arg	Ile	Lys	Leu	Thr	Gln	Leu	
caa	cgg	gcc	aga	tcc	caa	ggc	gta	ttc	ttt	gga	ggg	976
Gln	Arg	Ala	Arg	Ser	Gln	Gly	Val	Phe	Phe	Gly	Gly	
gga	gat	caa	cag	caa	ggg	gga	cta	ccc	att	ggc	cct	1024
Gly	Asp	Gln	Gln	Gln	Gly	Gly	Leu	Pro	Ile	Gly	Pro	
tct	gaa	gca	gcg	gtg	ttc	gat	atg	gaa	tat	gcg	agg	1072
Ser	Glu	Ala	Ala	Val	Phe	Asp	Met	Glu	Tyr	Ala	Arg	
cag	cag	agg	cta	tta	aac	gaa	cta	agg	gtg	gca	aca	1120
Gln	Gln	Arg	Leu	Leu	Asn	Glu	Leu	Arg	Val	Ala	Thr	
tcc	gag	aac	gag	ctt	agg	atg	ttt	gtg	gac	aca	tgt	1168
Ser	Glu	Asn	Glu	Leu	Arg	Met	Phe	Val	Asp	Thr	Cys	

## MBI0018 Sequence Listing.ST25

270	275	280	
gac cat ttg att aac ctc aag gct atg gtc gct aag acc gat gtc ttc			1216
Asp His Leu Ile Asn Leu Lys Ala Met Val Ala Lys Thr Asp Val Phe			
285	290	295	
cac ctc att tct gga gca tgg aaa act cca gct gaa cgt tgc ttc ttg			1264
His Leu Ile Ser Gly Ala Trp Lys Thr Pro Ala Glu Arg Cys Phe Leu			
300	305	310	
tgg atg ggt ggt ttc cgt cca tcg gag atc att aag gtg att gtg aac			1312
Trp Met Gly Gly Phe Arg Pro Ser Glu Ile Ile Lys Val Ile Val Asn			
315	320	325	
cag ata gaa cca ttg acg gag caa cag ata gtt ggg ata tgt ggg ctg			1360
Gln Ile Glu Pro Leu Thr Glu Gln Gln Ile Val Gly Ile Cys Gly Leu			
330	335	340	345
caa cag tcc aca caa gag gcc gag gag gct ctc tcg caa ggc ctc gag			1408
Gln Gln Ser Thr Gln Glu Ala Glu Glu Ala Leu Ser Gln Gly Leu Glu			
350	355	360	
gcg ttg aat caa tca ctt tcc gat agc att gtc tct gac tcc ctc ccg			1456
Ala Leu Asn Gln Ser Leu Ser Asp Ser Ile Val Ser Asp Ser Leu Pro			
365	370	375	
cct gcc tcc gca cca ctt cct cct cat cta tcc aat ttc atg tca cac			1504
Pro Ala Ser Ala Pro Leu Pro Pro His Leu Ser Asn Phe Met Ser His			
380	385	390	
atg tcc tta gct ctc aac aag ctc tct gct ctc gag ggc ttc gtt ctc			1552
Met Ser Leu Ala Leu Asn Lys Leu Ser Ala Leu Glu Gly Phe Val Leu			
395	400	405	
cag gcg gat aat ttg agg cac caa acg atc cat agg ctg aac caa ttg			1600
Gln Ala Asp Asn Leu Arg His Gln Thr Ile His Arg Leu Asn Gln Leu			
410	415	420	425
ttg acg acc cgt caa gaa gca cgg tgt ctt cta gcc gtt gcg gag tac			1648
Leu Thr Thr Arg Gln Glu Ala Arg Cys Leu Leu Ala Val Ala Glu Tyr			
430	435	440	
ttc cac cgt ctt caa gct cta agt tct ctc tgg cta gcc cgt cct cgg			1696
Phe His Arg Leu Gln Ala Leu Ser Ser Leu Trp Leu Ala Arg Pro Arg			
445	450	455	
caa gat gga taa tactaaaaca actgatgaag gaaacaaaa acaaaaacaa			1748
Gln Asp Gly			
460			
gagaataggt tgattagtta gccgccagct tgacctcttt atcatatata tcgtctctct			1808
actcaaatac agtgcaatta gggaaaattg tttggcttct ttttggtata tgattcttac			1868
tattatgttt ttaatcaaga			1888
<210> 18			
<211> 460			
<212> PRT			
<213> Arabidopsis thaliana			
<400> 18			
Met Gln Gly His His Gln Asn His His Gln His Leu Ser Ser Ser Ser			
1	5	10	15
Ala Thr Ser Ser His Gly Asn Phe Met Asn Lys Asp Gly Tyr Asp Ile			
20	25	30	
Gly Glu Ile Asp Pro Ser Leu Phe Leu Tyr Leu Asp Gly Gln Gly His			
35	40	45	

## MBI0018 Sequence Listing.ST25

His Asp Pro Pro Ser Thr Ala Pro Ser Pro Leu His His His His Thr  
 50 55 60  
 Thr Gln Asn Leu Ala Met Arg Pro Pro Thr Ser Thr Leu Asn Ile Phe  
 65 70 75 80  
 Pro Ser Gln Pro Met His Ile Glu Pro Pro Ser Ser Thr His Asn  
 85 90 95  
 Thr Asp Asn Thr Arg Leu Val Pro Ala Ala Gln Pro Ser Gly Ser Thr  
 100 105 110  
 Arg Pro Ala Ser Asp Pro Ser Met Asp Leu Thr Asn His Ser Gln Phe  
 115 120 125  
 His Gln Pro Pro Gln Gly Ser Lys Ser Ile Lys Lys Glu Gly Asn Arg  
 130 135 140  
 Lys Gly Leu Ala Ser Ser Asp His Asp Ile Pro Lys Ser Ser Asp Pro  
 145 150 155 160  
 Lys Thr Leu Arg Arg Leu Ala Gln Asn Arg Glu Ala Ala Arg Lys Ser  
 165 170 175  
 Arg Leu Arg Lys Lys Ala Tyr Val Gln Gln Leu Glu Ser Cys Arg Ile  
 180 185 190  
 Lys Leu Thr Gln Leu Glu Gln Glu Ile Gln Arg Ala Arg Ser Gln Gly  
 195 200 205  
 Val Phe Phe Gly Gly Ser Leu Ile Gly Gly Asp Gln Gln Gln Gly Gly  
 210 215 220  
 Leu Pro Ile Gly Pro Gly Asn Ile Ser Ser Glu Ala Ala Val Phe Asp  
 225 230 235 240  
 Met Glu Tyr Ala Arg Trp Leu Glu Glu Gln Gln Arg Leu Leu Asn Glu  
 245 250 255  
 Leu Arg Val Ala Thr Gln Glu His Leu Ser Glu Asn Glu Leu Arg Met  
 260 265 270  
 Phe Val Asp Thr Cys Leu Ala His Tyr Asp His Leu Ile Asn Leu Lys  
 275 280 285  
 Ala Met Val Ala Lys Thr Asp Val Phe His Leu Ile Ser Gly Ala Trp  
 290 295 300  
 Lys Thr Pro Ala Glu Arg Cys Phe Leu Trp Met Gly Gly Phe Arg Pro  
 305 310 315 320  
 Ser Glu Ile Ile Lys Val Ile Val Asn Gln Ile Glu Pro Leu Thr Glu  
 325 330 335  
 Gln Gln Ile Val Gly Ile Cys Gly Leu Gln Gln Ser Thr Gln Glu Ala  
 340 345 350

## MBI0018 Sequence Listing.ST25

Glu Glu Ala Leu Ser Gln Gly Leu Glu Ala Leu Asn Gln Ser Leu Ser  
 355 360 365

Asp Ser Ile Val Ser Asp Ser Leu Pro Pro Ala Ser Ala Pro Leu Pro  
 370 375 380

Pro His Leu Ser Asn Phe Met Ser His Met Ser Leu Ala Leu Asn Lys  
 385 390 395 400

Leu Ser Ala Leu Glu Gly Phe Val Leu Gln Ala Asp Asn Leu Arg His  
 405 410 415

Gln Thr Ile His Arg Leu Asn Gln Leu Leu Thr Thr Arg Gln Glu Ala  
 420 425 430

Arg Cys Leu Leu Ala Val Ala Glu Tyr Phe His Arg Leu Gln Ala Leu  
 435 440 445

Ser Ser Leu Trp Leu Ala Arg Pro Arg Gln Asp Gly  
 450 455 460

<210> 19  
 <211> 1707  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (98)..(1444)  
 <223> G748

<400> 19  
 ccacgcgtcc gcactctccc aaatctctct tctttaacaa caaaaaaaaa atcacagaga 60

catagagaga agaagacgga acagaggctc caaaaaa atg atg atg gag act aga 115  
 Met Met Met Glu Thr Arg  
 1 5

gat cca gct att aag ctt ttc ggt atg aaa atc cct ttt ccg tcg gtt 163  
 Asp Pro Ala Ile Lys Leu Phe Gly Met Lys Ile Pro Phe Pro Ser Val  
 10 15 20

ttt gaa tcg gca gtt acg gtg gag gat gac gaa gaa gat gac tgg agc 211  
 Phe Glu Ser Ala Val Thr Val Glu Asp Asp Glu Glu Asp Asp Trp Ser  
 25 30 35

ggc gga gat gac aaa tca cca gag aag gta act cca gag tta tca gat 259  
 Gly Gly Asp Asp Lys Ser Pro Glu Lys Val Thr Pro Glu Leu Ser Asp  
 40 45 50

aag aac aac aac aac tgt aac gac aac agt ttt aac aat tcg aaa ccc 307  
 Lys Asn Asn Asn Asn Cys Asn Asp Asn Ser Phe Asn Asn Ser Lys Pro  
 55 60 65 70

gaa acc ttg gac aaa gag gaa gcg aca tca act gat cag ata gag agt 355  
 Glu Thr Leu Asp Lys Glu Glu Ala Thr Ser Thr Asp Gln Ile Glu Ser  
 75 80 85

agt gac acg cct gag gat aat cag cag acg aca cct gat ggt aaa acc 403  
 Ser Asp Thr Pro Glu Asp Asn Gln Gln Thr Thr Pro Asp Gly Lys Thr  
 90 95 100

cta aag aaa ccg act aag att cta ccg tgt ccg aga tgc aaa agc atg 451  
 Leu Lys Lys Pro Thr Lys Ile Leu Pro Cys Pro Arg Cys Lys Ser Met  
 105 110 115



## MBI0018 Sequence Listing.ST25

gag acc aag ttc tgt tat tac aac aac tac aac ata aac cag cct cgt Glu Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr Asn Ile Asn Gln Pro Arg 120 125 130	499
cat ttc tgc aag gct tgt cag aga tat tgg act gct gga ggg act atg His Phe Cys Lys Ala Cys Gln Arg Tyr Trp Thr Ala Gly Gly Thr Met 135 140 145 150	547
agg aat gtt cct gtg ggg gca gga cgt cgt aag aac aaa agc tca tct Arg Asn Val Pro Val Gly Ala Gly Arg Lys Asn Lys Ser Ser Ser 155 160 165	595
tct cat tac cgt cac atc act att tcc gag gct ctt gag gct gcg agg Ser His Tyr Arg His Ile Thr Ile Ser Glu Ala Leu Glu Ala Arg 170 175 180	643
ctt gac ccg ggc tta cag gca aac aca agg gtc ttg agt ttt ggt ctc Leu Asp Pro Gly Leu Gln Ala Asn Thr Arg Val Leu Ser Phe Gly Leu 185 190 195	691
gaa gct cag cag cag cac gtt gct gct ccc atg aca cct gtt atg aag Glu Ala Gln Gln Gln His Val Ala Ala Pro Met Thr Pro Val Met Lys 200 205 210	739
cta caa gaa gat caa aag gtc tca aac ggt gct agg aac agg ttt cac Leu Gln Glu Asp Gln Lys Val Ser Asn Gly Ala Arg Asn Arg Phe His 215 220 225 230	787
ggg tta gcg gat caa cgg ctt gta gct cgg gta gag aat gga gat gat Gly Leu Ala Asp Gln Arg Leu Val Ala Arg Val Glu Asn Gly Asp Asp 235 240 245	835
tgc tca agc gga tcc tct gtg acc acc tct aac aat cac tca gtg gat Cys Ser Ser Gly Ser Ser Val Thr Thr Ser Asn Asn His Ser Val Asp 250 255 260	883
gaa tca aga gca caa agc ggc agt gtt gtt gaa gca caa atg aac aac Glu Ser Arg Ala Gln Ser Gly Ser Val Val Glu Ala Gln Met Asn Asn 265 270 275	931
aac aac aac aat aac atg aat ggt tat gct tgc atc cca ggt gtt cca Asn Asn Asn Asn Asn Met Asn Gly Tyr Ala Cys Ile Pro Gly Val Pro 280 285 290	979
tgg cct tac acg tgg aat cca gcg atg cct cca cca ggt ttt tac ccg Trp Pro Tyr Thr Trp Asn Pro Ala Met Pro Pro Pro Gly Phe Tyr Pro 295 300 305 310	1027
cct cca ggg tat cca atg ccg ttt tac cct tac tgg acc atc cca atg Pro Pro Gly Tyr Pro Met Pro Phe Tyr Pro Tyr Trp Thr Ile Pro Met 315 320 325	1075
cta cca ccg cat caa tcc tca tcg cct ata agc caa aag tgt tca aat Leu Pro Pro His Gln Ser Ser Ser Pro Ile Ser Gln Lys Cys Ser Asn 330 335 340	1123
aca aac tct ccg act ctc gga aag cat ccg aga gat gaa gga tca tcg Thr Asn Ser Pro Thr Leu Gly Lys His Pro Arg Asp Glu Gly Ser Ser 345 350 355	1171
aaa aag gac aat gag aca gag cga aaa cag aag gcc ggg tgc gtt ctg Lys Lys Asp Asn Glu Thr Glu Arg Lys Gln Lys Ala Gly Cys Val Leu 360 365 370	1219
gtc ccg aaa acg ttg aga ata gat gat cct aac gaa gca gca aag agc Val Pro Lys Thr Leu Arg Ile Asp Asp Pro Asn Glu Ala Ala Lys Ser 375 380 385 390	1267
tcg ata tgg aca aca ttg gga atc aag aac gag gcg atg tgc aaa gcc Ser Ile Trp Thr Thr Leu Gly Ile Lys Asn Glu Ala Met Cys Lys Ala 395 400 405	1315
ggt ggt atg ttc aaa ggg ttt gat cat aag aca aag atg tat aac aac Gly Gly Met Phe Lys Gly Phe Asp His Lys Thr Lys Met Tyr Asn Asn 410 415 420	1363

## MBI0018 Sequence Listing.ST25

```

410                               415                               420
gac aaa gct gag aac tcc cct gtt ctt tct gct aac cct gct gct cta 1411
Asp Lys Ala Glu Asn Ser Pro Val Leu Ser Ala Asn Pro Ala Ala Leu
425                               430                               435

tca aga tca cac aat ttc cat gaa cag att tag agttacatat gtatatgtat 1464
Ser Arg Ser His Asn Phe His Glu Gln Ile
440                               445

atatgtatga ttgattgtat gtatagatga tactggagaa tgatgagttt ttgagaatca 1524
aactcttttc ttctttctag tgattgcctt tattccttta catgttttgg ttctctgtac 1584
actatttgat ttaccttttt tactttcttt cttcatttgt caggaaatgt tggaagataa 1644
cattaatggg aaaaagttgg tgtggaccgt tgttgcgttg gcatttcaaa aaaaaaaaaa 1704
aaa 1707

<210> 20
<211> 448
<212> PRT
<213> Arabidopsis thaliana

<400> 20
Met Met Met Glu Thr Arg Asp Pro Ala Ile Lys Leu Phe Gly Met Lys
1 5 10 15

Ile Pro Phe Pro Ser Val Phe Glu Ser Ala Val Thr Val Glu Asp Asp
20 25 30

Glu Glu Asp Asp Trp Ser Gly Gly Asp Asp Lys Ser Pro Glu Lys Val
35 40 45

Thr Pro Glu Leu Ser Asp Lys Asn Asn Asn Asn Cys Asn Asp Asn Ser
50 55 60

Phe Asn Asn Ser Lys Pro Glu Thr Leu Asp Lys Glu Glu Ala Thr Ser
65 70 75 80

Thr Asp Gln Ile Glu Ser Ser Asp Thr Pro Glu Asp Asn Gln Gln Thr
85 90 95

Thr Pro Asp Gly Lys Thr Leu Lys Lys Pro Thr Lys Ile Leu Pro Cys
100 105 110

Pro Arg Cys Lys Ser Met Glu Thr Lys Phe Cys Tyr Tyr Asn Asn Tyr
115 120 125

Asn Ile Asn Gln Pro Arg His Phe Cys Lys Ala Cys Gln Arg Tyr Trp
130 135 140

Thr Ala Gly Gly Thr Met Arg Asn Val Pro Val Gly Ala Gly Arg Arg
145 150 155 160

Lys Asn Lys Ser Ser Ser Ser His Tyr Arg His Ile Thr Ile Ser Glu
165 170 175

Ala Leu Glu Ala Ala Arg Leu Asp Pro Gly Leu Gln Ala Asn Thr Arg
180 185 190

```

## MBI0018 Sequence Listing.ST25

Val Leu Ser Phe Gly Leu Glu Ala Gln Gln Gln His Val Ala Ala Pro  
 195 200 205  
 Met Thr Pro Val Met Lys Leu Gln Glu Asp Gln Lys Val Ser Asn Gly  
 210 215 220  
 Ala Arg Asn Arg Phe His Gly Leu Ala Asp Gln Arg Leu Val Ala Arg  
 225 230 235 240  
 Val Glu Asn Gly Asp Asp Cys Ser Ser Gly Ser Ser Val Thr Thr Ser  
 245 250 255  
 Asn Asn His Ser Val Asp Glu Ser Arg Ala Gln Ser Gly Ser Val Val  
 260 265 270  
 Glu Ala Gln Met Asn Asn Asn Asn Asn Asn Met Asn Gly Tyr Ala  
 275 280 285  
 Cys Ile Pro Gly Val Pro Trp Pro Tyr Thr Trp Asn Pro Ala Met Pro  
 290 295 300  
 Pro Pro Gly Phe Tyr Pro Pro Pro Gly Tyr Pro Met Pro Phe Tyr Pro  
 305 310 315 320  
 Tyr Trp Thr Ile Pro Met Leu Pro Pro His Gln Ser Ser Ser Pro Ile  
 325 330 335  
 Ser Gln Lys Cys Ser Asn Thr Asn Ser Pro Thr Leu Gly Lys His Pro  
 340 345 350  
 Arg Asp Glu Gly Ser Ser Lys Lys Asp Asn Glu Thr Glu Arg Lys Gln  
 355 360 365  
 Lys Ala Gly Cys Val Leu Val Pro Lys Thr Leu Arg Ile Asp Asp Pro  
 370 375 380  
 Asn Glu Ala Ala Lys Ser Ser Ile Trp Thr Thr Leu Gly Ile Lys Asn  
 385 390 395 400  
 Glu Ala Met Cys Lys Ala Gly Gly Met Phe Lys Gly Phe Asp His Lys  
 405 410 415  
 Thr Lys Met Tyr Asn Asn Asp Lys Ala Glu Asn Ser Pro Val Leu Ser  
 420 425 430  
 Ala Asn Pro Ala Ala Leu Ser Arg Ser His Asn Phe His Glu Gln Ile  
 435 440 445

<210> 21  
 <211> 1149  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)..(1149)  
 <223> G431

## MBI0018 Sequence Listing.ST25

```

<400> 21
atg gag agt ggt tcc aac agc act tct tgt cca atg gct ttt gcc ggg      48
Met Glu Ser Gly Ser Asn Ser Thr Ser Cys Pro Met Ala Phe Ala Gly
1 5 10 15

gat aat agt gat ggt ccg atg tgt cct atg atg atg atg atg ccg ccc      96
Asp Asn Ser Asp Gly Pro Met Cys Pro Met Met Met Met Met Pro Pro
20 25 30

atc atg aca tca cat caa cat cat ggt cat gat cat caa cat caa caa      144
Ile Met Thr Ser His Gln His His Gly His Asp His Gln His Gln Gln
35 40 45

caa gaa cat gat ggt tat gca tat cag tca cac cac caa caa agt agt      192
Gln Glu His Asp Gly Tyr Ala Tyr Gln Ser His His Gln Gln Ser Ser
50 55 60

tcc ctt ttt ctt caa tca cta gct cct ccc caa gga act aag aac aaa      240
Ser Leu Phe Leu Gln Ser Leu Ala Pro Pro Gln Gly Thr Lys Asn Lys
65 70 75 80

gtt gct tct tct tct tct cct tcc tct tgt gct cct gcc tat tct cta      288
Val Ala Ser Ser Ser Ser Pro Ser Ser Cys Ala Pro Ala Tyr Ser Leu
85 90 95

atg gag atc cat cat aac gaa atc gtt gca gga gga atc aac cct tgc      336
Met Glu Ile His His Asn Glu Ile Val Ala Gly Gly Ile Asn Pro Cys
100 105 110

tcc tct ttc tct tct tca gcc tct gtc aag gcc aag atc atg gct cat      384
Ser Ser Phe Ser Ser Ser Ala Ser Val Lys Ala Lys Ile Met Ala His
115 120 125

cct cac tac cac cgc ctc ttg gcc gct tat gtc aat tgt cag aag gtt      432
Pro His Tyr His Arg Leu Leu Ala Ala Tyr Val Asn Cys Gln Lys Val
130 135 140

gga gca cca ccg gag gtt gtg gcg agg ctg gag gag gca tgc tcg tct      480
Gly Ala Pro Pro Glu Val Val Ala Arg Leu Glu Glu Ala Cys Ser Ser
145 150 155 160

gcc gca gcc gca gcc gca tct atg ggg cca aca ggg tgt ctt ggt gaa      528
Ala Ala Ala Ala Ala Ala Ser Met Gly Pro Thr Gly Cys Leu Gly Glu
165 170 175

gat cca ggg ctt gat caa ttc atg gaa gct tac tgt gaa atg ctc gtt      576
Asp Pro Gly Leu Asp Gln Phe Met Glu Ala Tyr Cys Glu Met Leu Val
180 185 190

aag tat gag caa gag ctc tcc aaa cct ttc aag gaa gct atg gtc ttc      624
Lys Tyr Glu Gln Glu Leu Ser Lys Pro Phe Lys Glu Ala Met Val Phe
195 200 205

ctt caa cgt gtc gag tgt caa ttc aaa tcc ctc tct cta tcc tca cct      672
Leu Gln Arg Val Glu Cys Gln Phe Lys Ser Leu Ser Leu Ser Ser Pro
210 215 220

tcc tct ttc tcc ggt tat gga gag aca gca att gat agg aac aat aat      720
Ser Ser Phe Ser Gly Tyr Gly Glu Thr Ala Ile Asp Arg Asn Asn Asn
225 230 235 240

ggg tca tcc gag gaa gaa gtc gat atg aac aat gaa ttt gta gat cca      768
Gly Ser Ser Glu Glu Glu Val Asp Met Asn Asn Glu Phe Val Asp Pro
245 250 255

caa gct gag gat aga gag ctt aaa gga cag ctc ttg cgc aag tac agt      816
Gln Ala Glu Asp Arg Glu Leu Lys Gly Gln Leu Leu Arg Lys Tyr Ser
260 265 270

ggt tac tta ggg agc ctc aag caa gag ttc atg aag aag agg aag aaa      864
Gly Tyr Leu Gly Ser Leu Lys Gln Glu Phe Met Lys Lys Arg Lys Lys
275 280 285

gga aag ctc cct aaa gaa gct cgt caa caa ctg ctt gat tgg tgg agc      912

```

MBI0018 Sequence Listing.ST25

Gly	Lys	Leu	Pro	Lys	Glu	Ala	Arg	Gln	Gln	Leu	Leu	Asp	Trp	Trp	Ser		
290						295					300						
cgt	cac	tac	aaa	tgg	cct	tac	cct	tcg	gag	caa	caa	aag	ctc	gcc	ctt	960	
Arg	His	Tyr	Lys	Trp	Pro	Tyr	Pro	Ser	Glu	Gln	Gln	Lys	Leu	Ala	Leu		
305					310					315					320		
gcg	gaa	tca	acg	ggg	ctg	gac	cag	aaa	cag	ata	aac	aat	tgg	ttc	ata	1008	
Ala	Glu	Ser	Thr	Gly	Leu	Asp	Gln	Lys	Gln	Ile	Asn	Asn	Trp	Phe	Ile		
				325					330					335			
aac	cag	agg	aaa	cgg	cat	tgg	aag	ccg	tcg	gag	gac	atg	cag	ttt	gta	1056	
Asn	Gln	Arg	Lys	Arg	His	Trp	Lys	Pro	Ser	Glu	Asp	Met	Gln	Phe	Val		
			340					345					350				
gta	atg	gac	gca	aca	cat	cct	cac	cat	tac	ttc	atg	gat	aat	gtc	ttg	1104	
Val	Met	Asp	Ala	Thr	His	Pro	His	His	Tyr	Phe	Met	Asp	Asn	Val	Leu		
		355					360					365					
gac	aat	cct	ttc	cca	atg	gat	cac	atc	tcc	tcc	acc	atg	ctt	tga		1149	
Asp	Asn	Pro	Phe	Pro	Met	Asp	His	Ile	Ser	Ser	Thr	Met	Leu				
	370					375					380						
<210> 22																	
<211> 382																	
<212> PRT																	
<213> Arabidopsis thaliana																	
<400> 22																	
Met	Glu	Ser	Gly	Ser	Asn	Ser	Thr	Ser	Cys	Pro	Met	Ala	Phe	Ala	Gly		
1				5					10					15			
Asp	Asn	Ser	Asp	Gly	Pro	Met	Cys	Pro	Met	Met	Met	Met	Met	Pro	Pro		
			20					25					30				
Ile	Met	Thr	Ser	His	Gln	His	His	Gly	His	Asp	His	Gln	His	Gln	Gln		
		35					40					45					
Gln	Glu	His	Asp	Gly	Tyr	Ala	Tyr	Gln	Ser	His	His	Gln	Gln	Ser	Ser		
		50				55					60						
Ser	Leu	Phe	Leu	Gln	Ser	Leu	Ala	Pro	Pro	Gln	Gly	Thr	Lys	Asn	Lys		
65				70						75				80			
Val	Ala	Ser	Ser	Ser	Ser	Pro	Ser	Ser	Cys	Ala	Pro	Ala	Tyr	Ser	Leu		
				85					90					95			
Met	Glu	Ile	His	His	Asn	Glu	Ile	Val	Ala	Gly	Gly	Ile	Asn	Pro	Cys		
			100					105					110				
Ser	Ser	Phe	Ser	Ser	Ser	Ala	Ser	Val	Lys	Ala	Lys	Ile	Met	Ala	His		
		115					120					125					
Pro	His	Tyr	His	Arg	Leu	Leu	Ala	Ala	Tyr	Val	Asn	Cys	Gln	Lys	Val		
		130				135					140						
Gly	Ala	Pro	Pro	Glu	Val	Val	Ala	Arg	Leu	Glu	Glu	Ala	Cys	Ser	Ser		
145				150						155				160			
Ala	Ala	Ala	Ala	Ala	Ala	Ser	Met	Gly	Pro	Thr	Gly	Cys	Leu	Gly	Glu		
				165					170					175			

## MBI0018 Sequence Listing.ST25

Asp Pro Gly Leu Asp Gln Phe Met Glu Ala Tyr Cys Glu Met Leu Val  
 180 185 190  
 Lys Tyr Glu Gln Glu Leu Ser Lys Pro Phe Lys Glu Ala Met Val Phe  
 195 200 205  
 Leu Gln Arg Val Glu Cys Gln Phe Lys Ser Leu Ser Leu Ser Ser Pro  
 210 215 220  
 Ser Ser Phe Ser Gly Tyr Gly Glu Thr Ala Ile Asp Arg Asn Asn Asn  
 225 230 235 240  
 Gly Ser Ser Glu Glu Glu Val Asp Met Asn Asn Glu Phe Val Asp Pro  
 245 250 255  
 Gln Ala Glu Asp Arg Glu Leu Lys Gly Gln Leu Leu Arg Lys Tyr Ser  
 260 265 270  
 Gly Tyr Leu Gly Ser Leu Lys Gln Glu Phe Met Lys Lys Arg Lys Lys  
 275 280 285  
 Gly Lys Leu Pro Lys Glu Ala Arg Gln Gln Leu Leu Asp Trp Trp Ser  
 290 295 300  
 Arg His Tyr Lys Trp Pro Tyr Pro Ser Glu Gln Gln Lys Leu Ala Leu  
 305 310 315 320  
 Ala Glu Ser Thr Gly Leu Asp Gln Lys Gln Ile Asn Asn Trp Phe Ile  
 325 330 335  
 Asn Gln Arg Lys Arg His Trp Lys Pro Ser Glu Asp Met Gln Phe Val  
 340 345 350  
 Val Met Asp Ala Thr His Pro His His Tyr Phe Met Asp Asn Val Leu  
 355 360 365  
 Asp Asn Pro Phe Pro Met Asp His Ile Ser Ser Thr Met Leu  
 370 375 380

<210> 23  
 <211> 1136  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (118)..(1074)  
 <223> G187

<400> 23  
 tagacctctt aggaaaaaaa cctaaaaacc taatcccca acctaaaagg cttatctcat 60  
 ctcttcttct ttgtcttctt tactcttttt ttacctctct cttcattggt cttcacc 117  
 atg tct aat gaa acc aga gat ctc tac aac tac caa tac cct tca tcg 165  
 Met Ser Asn Glu Thr Arg Asp Leu Tyr Asn Tyr Gln Tyr Pro Ser Ser  
 1 5 10 15  
 ttt tcg ttg cac gaa atg atg aat ctg cct act tca aat cca tct tct 213  
 Phe Ser Leu His Glu Met Met Asn Leu Pro Thr Ser Asn Pro Ser Ser  
 20 25 30

## MBI0018 Sequence Listing.ST25

tat gga aac ctc cca tca caa aac ggt ttt aat cca tct act tat tcc Tyr Gly Asn Leu Pro Ser Gln Asn Gly Phe Asn Pro Ser Thr Tyr Ser 35 40 45	261
ttc acc gat tgt ctc caa agt tct cca gca gcg tat gaa tct cta ctt Phe Thr Asp Cys Leu Gln Ser Ser Pro Ala Ala Tyr Glu Ser Leu Leu 50 55 60	309
cag aaa act ttt ggt ctt tct ccc tct tcc tca gag gtt ttc aat tct Gln Lys Thr Phe Gly Leu Ser Pro Ser Ser Ser Glu Val Phe Asn Ser 65 70 75 80	357
tcg atc gat caa gaa ccg aac cgt gat gtt act aat gac gta atc aat Ser Ile Asp Gln Glu Pro Asn Arg Asp Val Thr Asn Asp Val Ile Asn 85 90 95	405
ggt ggt gca tgc aac gag act gaa act agg gtt tct cct tct aat tct Gly Gly Ala Cys Asn Glu Thr Glu Thr Arg Val Ser Pro Ser Asn Ser 100 105 110	453
tcc tct agt gag gct gat cac ccc ggt gaa gat tcc ggt aag agc cgg Ser Ser Ser Glu Ala Asp His Pro Gly Glu Asp Ser Gly Lys Ser Arg 115 120 125	501
agg aaa cga gag tta gtc ggt gaa gaa gat caa att tcc aaa aaa gtt Arg Lys Arg Glu Leu Val Gly Glu Glu Asp Gln Ile Ser Lys Lys Val 130 135 140	549
ggg aaa acg aaa aag act gag gtg aag aaa caa aga gag cca cga gtc Gly Lys Thr Lys Lys Thr Glu Val Lys Lys Lys Gln Arg Glu Pro Arg Val 145 150 155 160	597
tcg ttt atg act aaa agt gaa gtt gat cat ctt gaa gat ggt tat aga Ser Phe Met Thr Lys Ser Glu Val Asp His Leu Glu Asp Gly Tyr Arg 165 170 175	645
tgg aga aaa tac ggc caa aag gct gta aaa aat agc cct tat cca agg Trp Arg Lys Tyr Gly Gln Lys Ala Val Lys Asn Ser Pro Tyr Pro Arg 180 185 190	693
agt tac tat aga tgt aca aca caa aag tgc aac gtg aag aaa cga gtg Ser Tyr Tyr Arg Cys Thr Thr Gln Lys Cys Asn Val Lys Lys Arg Val 195 200 205	741
gag aga tcg ttc caa gat cca acg gtt gtg att aca act tac gag ggt Glu Arg Ser Phe Gln Asp Pro Thr Val Val Ile Thr Thr Tyr Glu Gly 210 215 220	789
caa cac aac cac ccg att ccg act aat ctt cga gga agt tct gcc gcg Gln His Asn His Pro Ile Pro Thr Asn Leu Arg Gly Ser Ser Ala Ala 225 230 235 240	837
gct gct atg ttc tcc gca gac ctc atg act cca aga agc ttt gca cat Ala Ala Met Phe Ser Ala Asp Leu Met Thr Pro Arg Ser Phe Ala His 245 250 255	885
gat atg ttt agg acg gca gct tat act aac ggc ggt tct gtg gcg gcg Asp Met Phe Arg Thr Ala Ala Tyr Thr Asn Gly Gly Ser Val Ala Ala 260 265 270	933
gct ttg gat tat gga tat gga caa agt ggt tat ggt agt gtg aat tca Ala Leu Asp Tyr Gly Tyr Gly Gln Ser Gly Tyr Gly Ser Val Asn Ser 275 280 285	981
aac cct agt tct cac caa gtg tat cat caa ggg ggt gag tat gag ctc Asn Pro Ser Ser His Gln Val Tyr His Gln Gly Gly Glu Tyr Glu Leu 290 295 300	1029
ttg agg gag att ttt cct tca att ttc ttt aag caa gag cct tga Leu Arg Glu Ile Phe Pro Ser Ile Phe Phe Lys Gln Glu Pro 305 310 315	1074
tcgatcattg ttataactac atatattata tatattgaga gagagaggta gagaaaaaaa	1134

## MBI0018 Sequence Listing.ST25

aa

1136

<210> 24  
 <211> 318  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 24

Met Ser Asn Glu Thr Arg Asp Leu Tyr Asn Tyr Gln Tyr Pro Ser Ser  
 1 5 10 15

Phe Ser Leu His Glu Met Met Asn Leu Pro Thr Ser Asn Pro Ser Ser  
 20 25 30

Tyr Gly Asn Leu Pro Ser Gln Asn Gly Phe Asn Pro Ser Thr Tyr Ser  
 35 40 45

Phe Thr Asp Cys Leu Gln Ser Ser Pro Ala Ala Tyr Glu Ser Leu Leu  
 50 55 60

Gln Lys Thr Phe Gly Leu Ser Pro Ser Ser Ser Glu Val Phe Asn Ser  
 65 70 75 80

Ser Ile Asp Gln Glu Pro Asn Arg Asp Val Thr Asn Asp Val Ile Asn  
 85 90 95

Gly Gly Ala Cys Asn Glu Thr Glu Thr Arg Val Ser Pro Ser Asn Ser  
 100 105 110

Ser Ser Ser Glu Ala Asp His Pro Gly Glu Asp Ser Gly Lys Ser Arg  
 115 120 125

Arg Lys Arg Glu Leu Val Gly Glu Glu Asp Gln Ile Ser Lys Lys Val  
 130 135 140

Gly Lys Thr Lys Lys Thr Glu Val Lys Lys Gln Arg Glu Pro Arg Val  
 145 150 155 160

Ser Phe Met Thr Lys Ser Glu Val Asp His Leu Glu Asp Gly Tyr Arg  
 165 170 175

Trp Arg Lys Tyr Gly Gln Lys Ala Val Lys Asn Ser Pro Tyr Pro Arg  
 180 185 190

Ser Tyr Tyr Arg Cys Thr Thr Gln Lys Cys Asn Val Lys Lys Arg Val  
 195 200 205

Glu Arg Ser Phe Gln Asp Pro Thr Val Val Ile Thr Thr Tyr Glu Gly  
 210 215 220

Gln His Asn His Pro Ile Pro Thr Asn Leu Arg Gly Ser Ser Ala Ala  
 225 230 235 240

Ala Ala Met Phe Ser Ala Asp Leu Met Thr Pro Arg Ser Phe Ala His  
 245 250 255

Asp Met Phe Arg Thr Ala Ala Tyr Thr Asn Gly Gly Ser Val Ala Ala



## MBI0018 Sequence Listing.ST25

260

265

270

Ala Leu Asp Tyr Gly Tyr Gly Gln Ser Gly Tyr Gly Ser Val Asn Ser  
275 280 285

Asn Pro Ser Ser His Gln Val Tyr His Gln Gly Gly Glu Tyr Glu Leu  
290 295 300

Leu Arg Glu Ile Phe Pro Ser Ile Phe Phe Lys Gln Glu Pro  
305 310 315

<210> 25  
<211> 2580  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (1)..(2580)  
<223> G470

<400> 25  
atg gcg agt tcg gag gtt tca atg aaa ggt aat cgt gga gga gat aac 48  
Met Ala Ser Ser Glu Val Ser Met Lys Gly Asn Arg Gly Gly Asp Asn  
1 5 10 15  
ttc tcc tcc tct ggt ttt agt gac cct aag gag act aga aat gtc tcc 96  
Phe Ser Ser Ser Gly Phe Ser Asp Pro Lys Glu Thr Arg Asn Val Ser  
20 25 30  
gtc gcc ggc gag ggg caa aaa agt aat tct acc cga tcc gct gcg gct 144  
Val Ala Gly Glu Gly Gln Lys Ser Asn Ser Thr Arg Ser Ala Ala Ala  
35 40 45  
gag cgt gct ttg gac cct gag gct gct ctt tac aga gag cta tgg cac 192  
Glu Arg Ala Leu Asp Pro Glu Ala Ala Leu Tyr Arg Glu Leu Trp His  
50 55 60  
gct tgt gct ggt ccg ctt gtg acg gtt cct aga caa gac gac cga gtc 240  
Ala Cys Ala Gly Pro Leu Val Thr Val Pro Arg Gln Asp Asp Arg Val  
65 70 75 80  
ttc tat ttt cct caa gga cac atc gag cag gtg gag gct tcg acg aac 288  
Phe Tyr Phe Pro Gln Gly His Ile Glu Gln Val Glu Ala Ser Thr Asn  
85 90 95  
cag gcg gca gaa caa cag atg cct ctc tat gat ctt ccg tca aag ctt 336  
Gln Ala Ala Glu Gln Gln Met Pro Leu Tyr Asp Leu Pro Ser Lys Leu  
100 105 110  
ctc tgt cga gtt att aat gta gat tta aag gca gag gca gat aca gat 384  
Leu Cys Arg Val Ile Asn Val Asp Leu Lys Ala Glu Ala Asp Thr Asp  
115 120 125  
gaa gtt tat gcg cag att act ctt ctt cct gag gct aat caa gac gag 432  
Glu Val Tyr Ala Gln Ile Thr Leu Leu Pro Glu Ala Asn Gln Asp Glu  
130 135 140  
aat gca att gag aaa gaa gcg cct ctt cct cca cct ccg agg ttc cag 480  
Asn Ala Ile Glu Lys Glu Ala Pro Leu Pro Pro Pro Pro Arg Phe Gln  
145 150 155 160  
gtg cat tcg ttc tgc aaa acc ttg act gca tcc gac aca agt aca cat 528  
Val His Ser Phe Cys Lys Thr Leu Thr Ala Ser Asp Thr Ser Thr His  
165 170 175  
ggt gga ttt tct gtt ctt agg cga cat gcg gat gaa tgt ctc cca cct 576  
Gly Gly Phe Ser Val Leu Arg Arg His Ala Asp Glu Cys Leu Pro Pro  
180 185 190

MBI0018 Sequence Listing.ST25																	
ctg Leu	gat Asp	atg Met 195	tct Ser	cga Arg	cag Gln	cct Pro	ccc Pro 200	act Thr	caa Gln	gag Glu	tta Leu	gtt Val 205	gca Ala	aag Lys	gat Asp	624	
ttg Leu	cat His 210	gca Ala	aat Asn	gag Glu	tgg Trp	cga Arg 215	ttc Phe	aga Arg	cat His	ata Ile	ttc Phe 220	cgg Arg	ggg Gly	caa Gln	cca Pro	672	
cgg Arg 225	agg Arg	cat His	ttg Leu	cta Leu	cag Gln 230	agt Ser	ggg Gly	tgg Trp	agt Ser	gtg Val 235	ttt Phe	gtt Val	agc Ser	tcc Ser	aaa Lys 240	720	
agg Arg	cta Leu	gtt Val	gca Ala	ggc Gly 245	gat Asp	gcg Ala	ttt Phe	ata Ile	ttt Phe 250	cta Leu	agg Arg	ggc Gly	gag Glu	aat Asn 255	gga Gly	768	
gaa Glu	tta Leu	aga Arg	gtt Val 260	ggt Gly	gta Val	agg Arg	cgt Arg	gcg Ala 265	atg Met	cga Arg	caa Gln	caa Gln	gga Gly 270	aac Asn	gtg Val	816	
ccg Pro	tct Ser	tct Ser 275	gtt Val	ata Ile	tct Ser	agc Ser	cat His 280	agc Ser	atg Met	cat His	ctt Leu	gga Gly 285	gta Val	ctg Leu	gcc Ala	864	
acc Thr	gca Ala 290	tgg Trp	cat His	gcc Ala	att Ile	tca Ser 295	aca Thr	ggg Gly	act Thr	atg Met	ttt Phe 300	aca Thr	gtc Val	tac Tyr	tac Tyr	912	
aaa Lys 305	ccc Pro	agg Arg	acg Thr	agc Ser	cca Pro 310	tct Ser	gag Glu	ttt Phe	att Ile	gtt Val 315	ccg Pro	ttc Phe	gat Asp	cag Gln	tat Tyr 320	960	
atg Met	gag Glu	tct Ser	gtt Val	aag Lys 325	aat Asn	aac Asn	tac Tyr	tct Ser	att Ile 330	ggc Gly	atg Met	aga Arg	ttc Phe	aaa Lys 335	atg Met	1008	
aga Arg	ttt Phe	gaa Glu	ggc Gly 340	gaa Glu	gag Glu	gct Ala	cct Pro	gag Glu 345	cag Gln	agg Arg	ttt Phe	act Thr	ggc Gly 350	aca Thr	atc Ile	1056	
gtt Val	ggg Gly	att Ile 355	gaa Glu	gag Glu	tct Ser	gat Asp	cct Pro 360	act Thr	agg Arg	tgg Trp	cca Pro	aaa Lys 365	tca Ser	aag Lys	tgg Trp	1104	
aga Arg	tcc Ser 370	ctc Leu	aag Lys	gtg Val	aga Arg	tgg Trp 375	gat Asp	gag Glu	act Thr	tct Ser	agt Ser 380	att Ile	cct Pro	cga Arg	cct Pro	1152	
gat Asp 385	aga Arg	gta Val	tct Ser	ccg Pro	tgg Trp 390	aaa Lys	gta Val	gag Glu	cca Pro	gct Ala 395	ctt Leu	gct Ala	cct Pro	cct Pro	gct Ala 400	1200	
ttg Leu	agt Ser	cct Pro	gtt Val	cca Pro 405	atg Met	cct Pro	agg Arg	cct Pro	aag Lys 410	agg Arg	ccc Pro	aga Arg	tca Ser	aat Asn 415	ata Ile	1248	
gca Ala	cct Pro	tca Ser	tct Ser 420	cct Pro	gac Asp	tct Ser	tcg Ser	atg Met 425	ctt Leu	acc Thr	aga Arg	gaa Glu 430	ggg Gly	aca Thr	act Thr	1296	
aag Lys	gca Ala	aac Asn 435	atg Met	gac Asp	cct Pro	tta Leu	cca Pro 440	gca Ala	agc Ser	gga Gly	ctt Leu	tca Ser 445	agg Arg	gtc Val	ttg Leu	1344	
caa Gln	ggg Gly 450	caa Gln	gaa Glu	tac Tyr	tcg Ser	acc Thr 455	ttg Leu	agg Arg	acg Thr	aaa Lys	cat His 460	act Thr	gag Glu	agt Ser	gta Val	1392	
gag Glu 465	tgt Cys	gat Asp	gct Ala	cct Pro	gag Glu 470	aat Asn	tct Ser	gtt Val	gtc Val	tgg Trp 475	caa Gln	tct Ser	tca Ser	gcg Ala	gat Asp 480	1440	
gat Asp	gat Asp	aag Lys	gtt Val	gac Asp 485	gtg Val	gtt Val	tcg Ser	ggg Gly	tct Ser 490	aga Arg	aga Arg	tat Tyr	gga Gly	tct Ser 495	gag Glu	1488	

## MBI0018 Sequence Listing.ST25

aac tgg atg tcc tca gcc agg cat gaa cct act tac aca gat ttg ctc Asn Trp Met Ser Ser Ala Arg His Glu Pro Thr Tyr Thr Asp Leu Leu 500 505 510	1536
tcc ggc ttt ggg act aac ata gat cca tcc cat ggt cag cgg ata cct Ser Gly Phe Gly Thr Asn Ile Asp Pro Ser His Gly Gln Arg Ile Pro 515 520 525	1584
ttt tat gac cat tca tca tca cct tct atg cct gca aag aga atc ttg Phe Tyr Asp His Ser Ser 535 Pro Ser Met Pro Ala Lys Arg Ile Leu 530 540	1632
agt gat tca gaa ggc aag ttc gat tat ctt gct aac cag tgg cag atg Ser Asp Ser Glu Gly Lys 550 Phe Asp Tyr Leu 555 Ala Asn Gln Trp Gln Met 545 560	1680
ata cac tct ggt ctc tcc ctg aag tta cat gaa tct cct aag gta cct Ile His Ser Gly Leu Ser Leu Lys Leu His Glu Ser Pro Lys Val Pro 565 570 575	1728
gca gca act gat gcg tct ctc caa ggg cga tgc aat gtt aaa tac agc Ala Ala Thr Asp Ala Ser Leu Gln Gly Arg Cys Asn Val Lys Tyr Ser 580 585 590	1776
gaa tat cct gtt ctt aat ggt cta tgc act gag aat gct ggt ggt aac Glu Tyr Pro Val Leu Asn Gly Leu Ser Thr Glu Asn Ala Gly Gly Asn 595 600 605	1824
tgg cca ata cgt cca cgt gct ttg aat tat tat gag gaa gtg gtc aat Trp Pro Ile Arg Pro Arg Ala Leu Asn Tyr Tyr Glu Glu Val Val Asn 610 615 620	1872
gct caa gcg caa gct cag gct agg gag caa gta aca aaa caa ccc ttc Ala Gln Ala Gln Ala Gln Ala Arg Glu Gln Val Thr Lys Gln Pro Phe 625 630 635 640	1920
acg ata caa gag gag aca gca aag tca aga gaa ggg aac tgc agg ctc Thr Ile Gln Glu Glu Thr Ala Lys Ser Arg Glu Gly Asn Cys Arg Leu 645 650 655	1968
ttt ggc att cct ctg acc aac aac atg aat ggg aca gac tca acc atg Phe Gly Ile Pro Leu Thr Asn Asn Met Asn Gly Thr Asp Ser Thr Met 660 665 670	2016
tct cag aga aac aac ttg aat gat gct gcg ggg ctt aca cag ata gca Ser Gln Arg Asn Asn Leu Asn Asp Ala Ala Gly Leu Thr Gln Ile Ala 675 680 685	2064
tca cca aag gtt cag gac ctt tca gat cag tca aaa ggg tca aaa tca Ser Pro Lys Val Gln Asp Leu Ser Asp Gln Ser Lys Gly Ser Lys Ser 690 695 700	2112
aca aac gat cat cgt gaa cag gga aga cca ttc cag act aat aat cct Thr Asn Asp His Arg Glu Gln Gly Arg Pro Phe Gln Thr Asn Asn Pro 705 710 715 720	2160
cat ccg aag gat gct caa acg aaa acc aac tca agt agg agt tgc aca His Pro Lys Asp Ala Gln Thr Lys Thr Asn Ser Ser Arg Ser Cys Thr 725 730 735	2208
aag gtt cac aag cag gga att gca ctt ggc cgt tca gtg gat ctt tca Lys Val His Lys Gln Gly Ile Ala Leu Gly Arg Ser Val Asp Leu Ser 740 745 750	2256
aag ttc caa aac tat gag gag tta gtc gct gag ctg gac agg ctg ttt Lys Phe Gln Asn Tyr Glu Glu Leu Val Ala Glu Leu Asp Arg Leu Phe 755 760 765	2304
gag ttc aat gga gag ttg atg gct cct aag aaa gat tgg ttg ata gtt Glu Phe Asn Gly Glu Leu Met Ala Pro Lys Lys Asp Trp Leu Ile Val 770 775 780	2352
tac aca gat gaa gag aat gat atg atg ctt gtt ggt gac gat cct tgg Tyr Thr Asp Glu Glu Asn Asp Met Met Leu Val Gly Asp Asp Pro Trp	2400

## MBI0018 Sequence Listing.ST25

785	790	795	800	
cag gag ttt tgt tgc atg gtt cgc aaa atc ttc ata tac acg aaa gag				2448
Gln Glu Phe Cys Cys Met Val Arg Lys Ile Phe Ile Tyr Thr Lys Glu	805	810	815	
gaa gtg agg aag atg aac ccg ggg act tta agc tgt agg agc gag gaa				2496
Glu Val Arg Lys Met Asn Pro Gly Thr Leu Ser Cys Arg Ser Glu Glu	820	825	830	
gaa gca gtt gtt ggg gaa gga tca gat gca aag gac gcc aag tct gca				2544
Glu Ala Val Val Gly Glu Gly Ser Asp Ala Lys Asp Ala Lys Ser Ala	835	840	845	
tca aat cct tca ttg tcc agc gct ggg aac tct taa				2580
Ser Asn Pro Ser Leu Ser Ser Ala Gly Asn Ser	850	855		

<210> 26  
 <211> 859  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 26

Met Ala Ser Ser Glu Val Ser Met Lys Gly Asn Arg Gly Gly Asp Asn	1	5	10	15
Phe Ser Ser Ser Gly Phe Ser Asp Pro Lys Glu Thr Arg Asn Val Ser	20	25	30	
Val Ala Gly Glu Gly Gln Lys Ser Asn Ser Thr Arg Ser Ala Ala Ala	35	40	45	
Glu Arg Ala Leu Asp Pro Glu Ala Ala Leu Tyr Arg Glu Leu Trp His	50	55	60	
Ala Cys Ala Gly Pro Leu Val Thr Val Pro Arg Gln Asp Asp Arg Val	65	70	75	80
Phe Tyr Phe Pro Gln Gly His Ile Glu Gln Val Glu Ala Ser Thr Asn	85	90	95	
Gln Ala Ala Glu Gln Gln Met Pro Leu Tyr Asp Leu Pro Ser Lys Leu	100	105	110	
Leu Cys Arg Val Ile Asn Val Asp Leu Lys Ala Glu Ala Asp Thr Asp	115	120	125	
Glu Val Tyr Ala Gln Ile Thr Leu Leu Pro Glu Ala Asn Gln Asp Glu	130	135	140	
Asn Ala Ile Glu Lys Glu Ala Pro Leu Pro Pro Pro Pro Arg Phe Gln	145	150	155	160
Val His Ser Phe Cys Lys Thr Leu Thr Ala Ser Asp Thr Ser Thr His	165	170	175	
Gly Gly Phe Ser Val Leu Arg Arg His Ala Asp Glu Cys Leu Pro Pro	180	185	190	
Leu Asp Met Ser Arg Gln Pro Pro Thr Gln Glu Leu Val Ala Lys Asp				

MBI0018 Sequence Listing.ST25  
200 205

```

195
Leu His Ala Asn Glu Trp Arg Phe Arg His Ile Phe Arg Gly Gln Pro
210 215 220
Arg Arg His Leu Leu Gln Ser Gly Trp Ser Val Phe Val Ser Ser Lys
225 230 235 240
Arg Leu Val Ala Gly Asp Ala Phe Ile Phe Leu Arg Gly Glu Asn Gly
245 250 255
Glu Leu Arg Val Gly Val Arg Arg Ala Met Arg Gln Gln Gly Asn Val
260 265 270
Pro Ser Ser Val Ile Ser Ser His Ser Met His Leu Gly Val Leu Ala
275 280 285
Thr Ala Trp His Ala Ile Ser Thr Gly Thr Met Phe Thr Val Tyr Tyr
290 295 300
Lys Pro Arg Thr Ser Pro Ser Glu Phe Ile Val Pro Phe Asp Gln Tyr
305 310 315 320
Met Glu Ser Val Lys Asn Asn Tyr Ser Ile Gly Met Arg Phe Lys Met
325 330 335
Arg Phe Glu Gly Glu Glu Ala Pro Glu Gln Arg Phe Thr Gly Thr Ile
340 345 350
Val Gly Ile Glu Glu Ser Asp Pro Thr Arg Trp Pro Lys Ser Lys Trp
355 360 365
Arg Ser Leu Lys Val Arg Trp Asp Glu Thr Ser Ser Ile Pro Arg Pro
370 375 380
Asp Arg Val Ser Pro Trp Lys Val Glu Pro Ala Leu Ala Pro Pro Ala
385 390 395 400
Leu Ser Pro Val Pro Met Pro Arg Pro Lys Arg Pro Arg Ser Asn Ile
405 410 415
Ala Pro Ser Ser Pro Asp Ser Ser Met Leu Thr Arg Glu Gly Thr Thr
420 425 430
Lys Ala Asn Met Asp Pro Leu Pro Ala Ser Gly Leu Ser Arg Val Leu
435 440 445
Gln Gly Gln Glu Tyr Ser Thr Leu Arg Thr Lys His Thr Glu Ser Val
450 455 460
Glu Cys Asp Ala Pro Glu Asn Ser Val Val Trp Gln Ser Ser Ala Asp
465 470 475 480
Asp Asp Lys Val Asp Val Val Ser Gly Ser Arg Arg Tyr Gly Ser Glu
485 490 495

```

MBI0018 Sequence Listing.ST25

Asn	Trp	Met	Ser	Ser	Ala	Arg	His	Glu	Pro	Thr	Tyr	Thr	Asp	Leu	Leu	500	505	510
Ser	Gly	Phe	Gly	Thr	Asn	Ile	Asp	Pro	Ser	His	Gly	Gln	Arg	Ile	Pro	515	520	525
Phe	Tyr	Asp	His	Ser	Ser	Ser	Pro	Ser	Met	Pro	Ala	Lys	Arg	Ile	Leu	530	535	540
Ser	Asp	Ser	Glu	Gly	Lys	Phe	Asp	Tyr	Leu	Ala	Asn	Gln	Trp	Gln	Met	545	550	555
Ile	His	Ser	Gly	Leu	Ser	Leu	Lys	Leu	His	Glu	Ser	Pro	Lys	Val	Pro	565	570	575
Ala	Ala	Thr	Asp	Ala	Ser	Leu	Gln	Gly	Arg	Cys	Asn	Val	Lys	Tyr	Ser	580	585	590
Glu	Tyr	Pro	Val	Leu	Asn	Gly	Leu	Ser	Thr	Glu	Asn	Ala	Gly	Gly	Asn	595	600	605
Trp	Pro	Ile	Arg	Pro	Arg	Ala	Leu	Asn	Tyr	Tyr	Glu	Glu	Val	Val	Asn	610	615	620
Ala	Gln	Ala	Gln	Ala	Gln	Ala	Arg	Glu	Gln	Val	Thr	Lys	Gln	Pro	Phe	625	630	635
Thr	Ile	Gln	Glu	Glu	Thr	Ala	Lys	Ser	Arg	Glu	Gly	Asn	Cys	Arg	Leu	645	650	655
Phe	Gly	Ile	Pro	Leu	Thr	Asn	Asn	Met	Asn	Gly	Thr	Asp	Ser	Thr	Met	660	665	670
Ser	Gln	Arg	Asn	Asn	Leu	Asn	Asp	Ala	Ala	Gly	Leu	Thr	Gln	Ile	Ala	675	680	685
Ser	Pro	Lys	Val	Gln	Asp	Leu	Ser	Asp	Gln	Ser	Lys	Gly	Ser	Lys	Ser	690	695	700
Thr	Asn	Asp	His	Arg	Glu	Gln	Gly	Arg	Pro	Phe	Gln	Thr	Asn	Asn	Pro	705	710	715
His	Pro	Lys	Asp	Ala	Gln	Thr	Lys	Thr	Asn	Ser	Ser	Arg	Ser	Cys	Thr	725	730	735
Lys	Val	His	Lys	Gln	Gly	Ile	Ala	Leu	Gly	Arg	Ser	Val	Asp	Leu	Ser	740	745	750
Lys	Phe	Gln	Asn	Tyr	Glu	Glu	Leu	Val	Ala	Glu	Leu	Asp	Arg	Leu	Phe	755	760	765
Glu	Phe	Asn	Gly	Glu	Leu	Met	Ala	Pro	Lys	Lys	Asp	Trp	Leu	Ile	Val	770	775	780
Tyr	Thr	Asp	Glu	Glu	Asn	Asp	Met	Met	Leu	Val	Gly	Asp	Asp	Pro	Trp	785	790	795

## MBI0018 Sequence Listing.ST25

Gln Glu Phe Cys Cys Met Val Arg Lys Ile Phe Ile Tyr Thr Lys Glu  
805 810 815

Glu Val Arg Lys Met Asn Pro Gly Thr Leu Ser Cys Arg Ser Glu Glu  
820 825 830

Glu Ala Val Val Gly Glu Gly Ser Asp Ala Lys Asp Ala Lys Ser Ala  
835 840 845

Ser Asn Pro Ser Leu Ser Ser Ala Gly Asn Ser  
850 855

<210> 27  
<211> 1519  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (197)..(1252)  
<223> G615

<400> 27  
ttttttcttt tctttctttt ttgtctggtg tgagaaattg tacgcttact atctctctct 60  
ctctctgccca gattctctct tttgatgat gtgaaagttg tgcttttgtt tcttaagaaa 120  
aaggcatatt ttttaactct gattcttggg tcttgattct tgattcttgg ttttttttag 180  
cttcttaagt tcggtg atg tcg tct tcc acc aat gac tac aac gat ggt aat 232  
Met Ser Ser Ser Thr Asn Asp Tyr Asn Asp Gly Asn  
1 5 10  
aac aat gga gtg tac cct ctc tct ctt tac ctt tct tca ctc tct ggc 280  
Asn Asn Gly Val Tyr Pro Leu Ser Leu Tyr Leu Ser Leu Ser Gly  
15 20 25  
cat caa gac atc att cat aat ccc tac aac cat cag tta aaa gca tct 328  
His Gln Asp Ile Ile His Asn Pro Tyr Asn His Gln Leu Lys Ala Ser  
30 35 40  
ccg ggc cat atg gta tca gca gtt cct gaa tct ctg atc gat tac atg 376  
Pro Gly His Met Val Ser Ala Val Pro Glu Ser Leu Ile Asp Tyr Met  
45 50 55 60  
gcg ttt aag tca aat aat gtt gtg aat caa caa ggc ttt gag ttt cct 424  
Ala Phe Lys Ser Asn Asn Val Val Asn Gln Gln Gly Phe Glu Phe Pro  
65 70 75  
gag gtg tca aag gaa atc aag aag gtg gtg aag aag gac cga cat agc 472  
Glu Val Ser Lys Glu Ile Lys Lys Val Val Lys Lys Asp Arg His Ser  
80 85 90  
aag att caa acg gca caa ggg att aga gac agg agg gtt agg ctt ttt 520  
Lys Ile Gln Thr Ala Gln Gly Ile Arg Asp Arg Arg Val Arg Leu Phe  
95 100 105  
att ggg att gct cgc caa ttc ttt gat ctt cag gat atg ttg ggg ttt 568  
Ile Gly Ile Ala Arg Gln Phe Phe Asp Leu Gln Asp Met Leu Gly Phe  
110 115 120  
gat aaa gct agt aaa acg tta gac tgg ctg ctc aag aag tca aga aaa 616  
Asp Lys Ala Ser Lys Thr Leu Asp Trp Leu Leu Lys Lys Ser Arg Lys  
125 130 135 140  
gcc atc aaa gag gtc gta caa gca aaa aac ctc aac aat gat gat gaa 664  
Ala Ile Lys Glu Val Val Gln Ala Lys Asn Leu Asn Asn Asp Asp Glu  
145 150 155

MBI0018 Sequence Listing.ST25

gat ttt gga aac att gga ggc gat gta gaa caa gaa gag gag aag gag Asp Phe Gly Asn Ile Gly Gly Asp Val Glu Gln Glu Glu Glu Lys Glu 160 165 170	712
gag gat gac aat ggc gat aag agc ttc gtg tat ggt ttg agc ccc ggg Glu Asp Asp Asn Gly Asp Lys Ser Phe Val Tyr Gly Leu Ser Pro Gly 175 180 185	760
tac ggt gaa gaa gaa gtg gta tgt gag gcc acg aag gca ggg ata aga Tyr Gly Glu Glu Glu Val Val Cys Glu Ala Thr Lys Ala Gly Ile Arg 190 195 200	808
aag aag aag agt gag ttg aga aac atc tca tca aag ggg cta gga gcc Lys Lys Lys Ser Glu Leu Arg Asn Ile Ser Ser Lys Gly Leu Gly Ala 205 210 215 220	856
aaa gct aga gga aaa gca aag gag cga aca aaa gag atg atg gcc tat Lys Ala Arg Gly Lys Ala Lys Glu Arg Thr Lys Glu Met Met Ala Tyr 225 230 235	904
gat aat cca gag act gcc tct gat att aca caa tct gaa atc atg gac Asp Asn Pro Glu Thr Ala Ser Asp Ile Thr Gln Ser Glu Ile Met Asp 240 245 250	952
cca ttc aag agg tct ata gtc ttc aat gaa gga gaa gat atg aca cac Pro Phe Lys Arg Ser Ile Val Phe Asn Glu Gly Glu Asp Met Thr His 255 260 265	1000
ctt ttc tac aag gaa cca atc gag gag ttt gat aat caa gaa tct atc Leu Phe Tyr Lys Glu Pro Ile Glu Glu Phe Asp Asn Gln Glu Ser Ile 270 275 280	1048
tta acc aat atg act cta cca acg aag atg ggt caa agt tac aat caa Leu Thr Asn Met Thr Leu Pro Thr Lys Met Gly Gln Ser Tyr Asn Gln 285 290 295 300	1096
aat aat ggg ata ctt atg ttg gta gat cag agt tct agc agc aac tat Asn Asn Gly Ile Leu Met Leu Val Asp Gln Ser Ser Ser Ser Asn Tyr 305 310 315	1144
aat aca ttt ctg cct caa aat ttg gat tat agt tat gat caa aac cct Asn Thr Phe Leu Pro Gln Asn Leu Asp Tyr Ser Tyr Asp Gln Asn Pro 320 325 330	1192
ttt cat gac caa acc tta tat gta gtc acc gac aaa aat ttc ccc aaa Phe His Asp Gln Thr Leu Tyr Val Val Thr Asp Lys Asn Phe Pro Lys 335 340 345	1240
ggt ttc cta taa atctcgacag ttttgaagga ctatgcatga tcaagtttaa Gly Phe Leu 350	1292
acatgtaagc caatatagtc ccttattcct ctgaatgtat acaaaatcta tagttatgta	1352
tatctgttcc tttttaacgt atctttattg atcttctgtg ccttgatcaa aattgtcatt	1412
ttaagattca gtttggtgtaa tatttttagct acaactttta agtggtatta ttgtaacctt	1472
ttgaactata tatttttgaag atgaataaga acatgtttat ataaaaa	1519
 <210> 28	
<211> 351	
<212> PRT	
<213> Arabidopsis thaliana	
 <400> 28	
Met Ser Ser Ser Thr Asn Asp Tyr Asn Asp Gly Asn Asn Asn Gly Val 1 5 10 15	
Tyr Pro Leu Ser Leu Tyr Leu Ser Ser Leu Ser Gly His Gln Asp Ile 20 25 30	



## MBI0018 Sequence Listing.ST25

```

Ile His Asn Pro Tyr Asn His Gln Leu Lys Ala Ser Pro Gly His Met
 35                                40                                45

Val Ser Ala Val Pro Glu Ser Leu Ile Asp Tyr Met Ala Phe Lys Ser
 50                                55                                60

Asn Asn Val Val Asn Gln Gln Gly Phe Glu Phe Pro Glu Val Ser Lys
 65                                70                                75                                80

Glu Ile Lys Lys Val Val Lys Lys Asp Arg His Ser Lys Ile Gln Thr
 85                                90                                95

Ala Gln Gly Ile Arg Asp Arg Arg Val Arg Leu Phe Ile Gly Ile Ala
100                                105                                110

Arg Gln Phe Phe Asp Leu Gln Asp Met Leu Gly Phe Asp Lys Ala Ser
115                                120                                125

Lys Thr Leu Asp Trp Leu Leu Lys Lys Ser Arg Lys Ala Ile Lys Glu
130                                135                                140

Val Val Gln Ala Lys Asn Leu Asn Asn Asp Asp Glu Asp Phe Gly Asn
145                                150                                155                                160

Ile Gly Gly Asp Val Glu Gln Glu Glu Glu Lys Glu Glu Asp Asp Asn
165                                170                                175

Gly Asp Lys Ser Phe Val Tyr Gly Leu Ser Pro Gly Tyr Gly Glu Glu
180                                185                                190

Glu Val Val Cys Glu Ala Thr Lys Ala Gly Ile Arg Lys Lys Lys Ser
195                                200                                205

Glu Leu Arg Asn Ile Ser Ser Lys Gly Leu Gly Ala Lys Ala Arg Gly
210                                215                                220

Lys Ala Lys Glu Arg Thr Lys Glu Met Met Ala Tyr Asp Asn Pro Glu
225                                230                                235                                240

Thr Ala Ser Asp Ile Thr Gln Ser Glu Ile Met Asp Pro Phe Lys Arg
245                                250                                255

Ser Ile Val Phe Asn Glu Gly Glu Asp Met Thr His Leu Phe Tyr Lys
260                                265                                270

Glu Pro Ile Glu Glu Phe Asp Asn Gln Glu Ser Ile Leu Thr Asn Met
275                                280                                285

Thr Leu Pro Thr Lys Met Gly Gln Ser Tyr Asn Gln Asn Asn Gly Ile
290                                295                                300

Leu Met Leu Val Asp Gln Ser Ser Ser Ser Asn Tyr Asn Thr Phe Leu
305                                310                                315                                320

Pro Gln Asn Leu Asp Tyr Ser Tyr Asp Gln Asn Pro Phe His Asp Gln
325                                330                                335

```

## MBI0018 Sequence Listing.ST25

Thr Leu Tyr Val Val Thr Asp Lys Asn Phe Pro Lys Gly Phe Leu  
 340 345 350

<210> 29  
 <211> 974  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (62)..(874)  
 <223> G1073

<400> 29  
 cccccgcacc tgcctctaca gagacctgaa gattccagaa cccacactga tcaaaaataa 60  
 c atg gaa ctt aac aga tct gaa gca gac gaa gca aag gcc gag acc act 109  
 Met Glu Leu Asn Arg Ser Glu Ala Asp Glu Ala Lys Ala Glu Thr Thr  
 1 5 10 15  
 ccc acc ggt gga gcc acc agc tca gcc aca gcc tct ggc tct tcc tcc 157  
 Pro Thr Gly Gly Ala Thr Ser Ser Ala Thr Ala Ser Gly Ser Ser Ser  
 20 25 30  
 gga cgt cgt cca cgt ggt cgt cct gca ggt tcc aaa aac aaa ccc aaa 205  
 Gly Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys  
 35 40 45  
 cct ccg acg att ata act aga gat agt cct aac gtc ctt aga tca cac 253  
 Pro Pro Thr Ile Ile Thr Arg Asp Ser Pro Asn Val Leu Arg Ser His  
 50 55 60  
 gtt ctt gaa gtc acc tcc ggt tcg gac ata tcc gag gca gtc tcc acc 301  
 Val Leu Glu Val Thr Ser Gly Ser Asp Ile Ser Glu Ala Val Ser Thr  
 65 70 75 80  
 tac gcc act cgt cgc ggc tgc ggc gtt tgc att ata agc ggc acg ggt 349  
 Tyr Ala Thr Arg Arg Gly Cys Gly Val Cys Ile Ile Ser Gly Thr Gly  
 85 90 95  
 gcg gtc act aac gtc acg ata cgg caa cct gcg gct ccg gct ggt gga 397  
 Ala Val Thr Asn Val Thr Ile Arg Gln Pro Ala Ala Pro Ala Gly Gly  
 100 105 110  
 ggt gtg att acc ctg cat ggt cgg ttt gac att ttg tct ttg acc ggt 445  
 Gly Val Ile Thr Leu His Gly Arg Phe Asp Ile Leu Ser Leu Thr Gly  
 115 120 125  
 act gcg ctt cca ccg cct gca cca ccg gga gca gga ggt ttg acg gtg 493  
 Thr Ala Leu Pro Pro Pro Ala Pro Pro Gly Ala Gly Gly Leu Thr Val  
 130 135 140  
 tat cta gcc gga ggt caa gga caa gtt gta gga ggg aat gtg gct ggt 541  
 Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Asn Val Ala Gly  
 145 150 155 160  
 tcg tta att gct tcg gga ccg gta gtg ttg atg gct gct tct ttt gca 589  
 Ser Leu Ile Ala Ser Gly Pro Val Val Leu Met Ala Ala Ser Phe Ala  
 165 170 175  
 aac gca gtt tat gat agg tta ccg att gaa gag gaa gaa acc cca ccg 637  
 Asn Ala Val Tyr Asp Arg Leu Pro Ile Glu Glu Glu Glu Thr Pro Pro  
 180 185 190  
 ccg aga acc acc ggg gtg cag cag cag cag ccg gag gcg tct cag tcg 685  
 Pro Arg Thr Thr Gly Val Gln Gln Gln Gln Pro Glu Ala Ser Gln Ser  
 195 200 205  
 tcg gag gtt acg ggg agt ggg gcc cag gcg tgt gag tca aac ctc caa 733  
 Ser Glu Val Thr Gly Ser Gly Ala Gln Ala Cys Glu Ser Asn Leu Gln  
 210 215 220

## MBI0018 Sequence Listing.ST25

```

ggt gga aat ggt gga gga ggt gtt gct ttc tac aat ctt gga atg aat      781
Gly Gly Asn Gly Gly Gly Gly Val Ala Phe Tyr Asn Leu Gly Met Asn
225                230                235                240

atg aac aat ttt caa ttc tcc ggg gga gat att tac ggt atg agc ggc      829
Met Asn Asn Phe Gln Phe Ser Gly Gly Asp Ile Tyr Gly Met Ser Gly
                245                250                255

ggt agc gga gga ggt ggt ggc ggt gcg act aga ccc gcg ttt tag      874
Gly Ser Gly Gly Gly Gly Gly Gly Ala Thr Arg Pro Ala Phe
                260                265                270

agtttttagcg ttttggtgac accttttggt gcgtttgcgt gtttgacctc aaactactag      934

gctactagct atagcggttg cgaaatgcga atattagggtt      974

```

```

<210> 30
<211> 270
<212> PRT
<213> Arabidopsis thaliana

```

```

<400> 30

```

```

Met Glu Leu Asn Arg Ser Glu Ala Asp Glu Ala Lys Ala Glu Thr Thr
1                5                10                15

```

```

Pro Thr Gly Gly Ala Thr Ser Ser Ala Thr Ala Ser Gly Ser Ser Ser
                20                25                30

```

```

Gly Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys
                35                40                45

```

```

Pro Pro Thr Ile Ile Thr Arg Asp Ser Pro Asn Val Leu Arg Ser His
50                55                60

```

```

Val Leu Glu Val Thr Ser Gly Ser Asp Ile Ser Glu Ala Val Ser Thr
65                70                75                80

```

```

Tyr Ala Thr Arg Arg Gly Cys Gly Val Cys Ile Ile Ser Gly Thr Gly
                85                90                95

```

```

Ala Val Thr Asn Val Thr Ile Arg Gln Pro Ala Ala Pro Ala Gly Gly
100                105                110

```

```

Gly Val Ile Thr Leu His Gly Arg Phe Asp Ile Leu Ser Leu Thr Gly
115                120                125

```

```

Thr Ala Leu Pro Pro Pro Ala Pro Pro Gly Ala Gly Gly Leu Thr Val
130                135                140

```

```

Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Asn Val Ala Gly
145                150                155                160

```

```

Ser Leu Ile Ala Ser Gly Pro Val Val Leu Met Ala Ala Ser Phe Ala
165                170                175

```

```

Asn Ala Val Tyr Asp Arg Leu Pro Ile Glu Glu Glu Glu Thr Pro Pro
180                185                190

```

```

Pro Arg Thr Thr Gly Val Gln Gln Gln Gln Pro Glu Ala Ser Gln Ser
195                200                205

```

## MBI0018 Sequence Listing.ST25

Ser Glu Val Thr Gly Ser Gly Ala Gln Ala Cys Glu Ser Asn Leu Gln  
210 215 220

Gly Gly Asn Gly Gly Gly Gly Val Ala Phe Tyr Asn Leu Gly Met Asn  
225 230 235 240

Met Asn Asn Phe Gln Phe Ser Gly Gly Asp Ile Tyr Gly Met Ser Gly  
245 250 255

Gly Ser Gly Gly Gly Gly Gly Gly Ala Thr Arg Pro Ala Phe  
260 265 270

<210> 31  
<211> 2010  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (1)..(2010)  
<223> G1493

<400> 31  
atg atg aat ccg agt cac gga aga gga ctc gga tcg gct ggt ggg tcc 48  
Met Met Asn Pro Ser His Gly Arg Gly Leu Gly Ser Ala Gly Gly Ser  
1 5 10 15  
agc tcc ggt aga aat caa gga ggt ggt ggt gag acc gtc gtc gag atg 96  
Ser Ser Gly Arg Asn Gln Gly Gly Gly Gly Glu Thr Val Val Glu Met  
20 25 30  
ttt cct tct ggt ctt cga gtt ctt gtc gtt gac gat gac cca act tgt 144  
Phe Pro Ser Gly Leu Arg Val Leu Val Val Asp Asp Asp Pro Thr Cys  
35 40 45  
ctc atg atc tta gag agg atg ctt agg act tgt ctt tac gaa gta acg 192  
Leu Met Ile Leu Glu Arg Met Leu Arg Thr Cys Leu Tyr Glu Val Thr  
50 55 60  
aaa tgc aac aga gca gag atg gca ttg tct ctg ctc cgg aag aac aaa 240  
Lys Cys Asn Arg Ala Glu Met Ala Leu Ser Leu Leu Arg Lys Asn Lys  
65 70 75 80  
cat gga ttc gat ata gta atc agt gat gtt cat atg cct gac atg gac 288  
His Gly Phe Asp Ile Val Ile Ser Asp Val His Met Pro Asp Met Asp  
85 90 95  
ggt ttc aag ctt ctt gag cat gtt ggt cta gag atg gac tta cct gtt 336  
Gly Phe Lys Leu Leu Glu His Val Gly Leu Glu Met Asp Leu Pro Val  
100 105 110  
atc atg atg tct gcg gat gat tca aag agt gtg gtt cta aag gga gta 384  
Ile Met Met Ser Ala Asp Asp Ser Lys Ser Val Val Leu Lys Gly Val  
115 120 125  
acg cac ggt gcg gtt gat tac ctt atc aag cct gta cgt atg gag gca 432  
Thr His Gly Ala Val Asp Tyr Leu Ile Lys Pro Val Arg Met Glu Ala  
130 135 140  
ctt aag aac ata tgg cag cat gta gtt agg aag agg aga agt gaa tgg 480  
Leu Lys Asn Ile Trp Gln His Val Val Arg Lys Arg Arg Ser Glu Trp  
145 150 155 160  
agt gta ccg gaa cat tct ggg agc att gag gag act ggc gag aga cag 528  
Ser Val Pro Glu His Ser Gly Ser Ile Glu Glu Thr Gly Glu Arg Gln  
165 170 175  
cag cag caa cat aga gga ggt ggt ggt ggt gca gct gtt tct ggt gga 576

MBI0018 Sequence Listing.ST25																	
Gln	Gln	Gln	His	Arg	Gly	Gly	Gly	Gly	Ala	Ala	Val	Ser	Gly	Gly			
180							185					190					
gag	gat	gcg	gtg	gat	gat	aac	tca	tcc	tcg	gtt	aac	gaa	ggt	aac	aat	624	
Glu	Asp	Ala	Val	Asp	Asp	Asn	Ser	Ser	Ser	Val	Asn	Glu	Gly	Asn	Asn		
		195					200					205					
tgg	agg	agc	agt	tca	cgg	aag	agg	aaa	gac	gag	gaa	gga	gaa	gag	caa	672	
Trp	Arg	Ser	Ser	Ser	Arg	Lys	Arg	Lys	Asp	Glu	Glu	Gly	Glu	Glu	Gln		
	210					215					220						
gga	gat	gat	aag	gac	gaa	gat	gcg	tcg	aat	ttg	aag	aaa	ccg	cgt	gtc	720	
Gly	Asp	Asp	Lys	Asp	Glu	Asp	Ala	Ser	Asn	Leu	Lys	Lys	Pro	Arg	Val		
	225				230					235					240		
gtc	tgg	tct	gtt	gaa	ttg	cat	cag	cag	ttt	gtt	gct	gct	gtt	aat	cag	768	
Val	Trp	Ser	Val	Glu	Leu	His	Gln	Gln	Phe	Val	Ala	Ala	Val	Asn	Gln		
			245						250					255			
ctc	ggc	gtt	gag	aag	gcg	gtt	cct	aaa	aag	atc	tta	gag	ctg	atg	aat	816	
Leu	Gly	Val	Glu	Lys	Ala	Val	Pro	Lys	Lys	Ile	Leu	Glu	Leu	Met	Asn		
			260					265					270				
gtt	cct	ggt	cta	acc	cga	gaa	aac	gta	gca	agt	cac	ctc	cag	aaa	tac	864	
Val	Pro	Gly	Leu	Thr	Arg	Glu	Asn	Val	Ala	Ser	His	Leu	Gln	Lys	Tyr		
		275					280					285					
cgg	ata	tat	cta	aga	cgg	ctt	gga	ggg	gta	tcg	cag	cac	caa	ggc	aat	912	
Arg	Ile	Tyr	Leu	Arg	Arg	Leu	Gly	Gly	Val	Ser	Gln	His	Gln	Gly	Asn		
	290					295					300						
ctt	aac	aac	tcg	ttt	atg	acg	ggt	cag	gat	gcg	agc	ttc	gga	cct	ctt	960	
Leu	Asn	Asn	Ser	Phe	Met	Thr	Gly	Gln	Asp	Ala	Ser	Phe	Gly	Pro	Leu		
	305				310					315					320		
tcg	aca	ttg	aat	ggg	ttt	gat	ctt	caa	gca	cta	gcc	gtc	aca	ggt	cag	1008	
Ser	Thr	Leu	Asn	Gly	Phe	Asp	Leu	Gln	Ala	Leu	Ala	Val	Thr	Gly	Gln		
			325						330					335			
tta	cct	gca	cag	agt	ctt	gca	cag	ctt	caa	gcc	gct	ggt	tta	ggc	cgg	1056	
Leu	Pro	Ala	Gln	Ser	Leu	Ala	Gln	Leu	Gln	Ala	Ala	Gly	Leu	Gly	Arg		
			340					345					350				
cct	gcg	atg	gtc	tct	aag	tca	ggt	ttg	ccg	gtt	tcc	tcc	att	gtg	gat	1104	
Pro	Ala	Met	Val	Ser	Lys	Ser	Gly	Leu	Pro	Val	Ser	Ser	Ile	Val	Asp		
		355					360					365					
gag	aga	agc	atc	ttc	agc	ttt	gac	aac	acg	aaa	aca	aga	ttt	gga	gaa	1152	
Glu	Arg	Ser	Ile	Phe	Ser	Phe	Asp	Asn	Thr	Lys	Thr	Arg	Phe	Gly	Glu		
	370					375					380						
ggg	ctt	ggg	cat	cac	ggg	caa	caa	ccc	caa	cag	caa	cca	cag	atg	aac	1200	
Gly	Leu	Gly	His	His	Gly	Gln	Gln	Pro	Gln	Gln	Gln	Pro	Gln	Met	Asn		
	385				390					395					400		
tta	ctt	cac	ggt	gtc	ccc	acg	ggt	tta	caa	cag	cag	ctt	cct	atg	ggt	1248	
Leu	Leu	His	Gly	Val	Pro	Thr	Gly	Leu	Gln	Gln	Gln	Leu	Pro	Met	Gly		
			405					410						415			
aat	cga	atg	agt	att	caa	caa	cag	att	gct	gct	gtt	cga	gct	gga	aat	1296	
Asn	Arg	Met	Ser	Ile	Gln	Gln	Gln	Ile	Ala	Ala	Val	Arg	Ala	Gly	Asn		
			420					425					430				
agt	gtt	caa	aac	aac	gga	atg	ctg	atg	cct	cta	gcg	ggt	cag	cag	tct	1344	
Ser	Val	Gln	Asn	Asn	Gly	Met	Leu	Met	Pro	Leu	Ala	Gly	Gln	Gln	Ser		
		435					440					445					
ttg	cct	cgg	gga	cca	ccg	cct	atg	cta	acc	tct	tcg	caa	tca	tcc	atc	1392	
Leu	Pro	Arg	Gly	Pro	Pro	Pro	Met	Leu	Thr	Ser	Ser	Gln	Ser	Ser	Ile		
		450				455					460						
agg	cag	ccg	atg	tta	tca	aac	cgc	att	tcc	gag	aga	agt	ggt	ttc	tct	1440	
Arg	Gln	Pro	Met	Leu	Ser	Asn	Arg	Ile	Ser	Glu	Arg	Ser	Gly	Phe	Ser		
					470					475					480		

MBI0018 Sequence Listing.ST25

gga agg aac aat atc ccc gag agc agc aga gtg tta ccg aca agt tac	1488
Gly Arg Asn Asn Ile Pro Glu Ser Ser Arg Val Leu Pro Thr Ser Tyr	
485 490 495	
act aat ctc aca aca caa cac tca tca agc tcg atg cct tat aac aac	1536
Thr Asn Leu Thr Thr Gln His Ser Ser Ser Ser Met Pro Tyr Asn Asn	
500 505 510	
ttc caa cca gaa ctt ccc gtg aac agt ttc ccg ctg gca agt gca cca	1584
Phe Gln Pro Glu Leu Pro Val Asn Ser Phe Pro Leu Ala Ser Ala Pro	
515 520 525	
ggg ata tca gta ccg gtt cgg aaa gcc act tct tac cag gaa gag gtt	1632
Gly Ile Ser Val Pro Val Arg Lys Ala Thr Ser Tyr Gln Glu Glu Val	
530 535 540	
aac agc tcc gaa gcg ggt ttc att acg ccg agc tac gac atg ttc acc	1680
Asn Ser Ser Glu Ala Gly Phe Ile Thr Pro Ser Tyr Asp Met Phe Thr	
545 550 555 560	
acc aga cag aat gat tgg gat ctg agg aat att gga ata gcc ttt gac	1728
Thr Arg Gln Asn Asp Trp Asp Leu Arg Asn Ile Gly Ile Ala Phe Asp	
565 570 575	
tca cat cag gac tca gaa tcc gct gcg ttt tcc gct tca gaa gcc tac	1776
Ser His Gln Asp Ser Glu Ser Ala Ala Phe Ser Ala Ser Glu Ala Tyr	
580 585 590	
tct tct tcg tcc atg tca aga cac aac acg aca gtt gca gcc acc gag	1824
Ser Ser Ser Ser Met Ser Arg His Asn Thr Thr Val Ala Ala Thr Glu	
595 600 605	
cat ggc cga aac cac cag cag cca cca tcg gga atg gta cag cac cat	1872
His Gly Arg Asn His Gln Gln Pro Pro Ser Gly Met Val Gln His His	
610 615 620	
cag gtt tat gca gac gga aac ggt ggt tca gtg agg gtg aaa tca gag	1920
Gln Val Tyr Ala Asp Gln Asn Gly Gly Ser Val Arg Val Lys Ser Glu	
625 630 635 640	
aga gtg gct acg gat aca gca aca atg gcg ttt cac gag cag tat agt	1968
Arg Val Ala Thr Asp Thr Ala Thr Met Ala Phe His Glu Gln Tyr Ser	
645 650 655	
aat caa gaa gat ctt atg agc gca ctt ctt aag cag gtt tga	2010
Asn Gln Glu Asp Leu Met Ser Ala Leu Leu Lys Gln Val	
660 665	
<210> 32	
<211> 669	
<212> PRT	
<213> Arabidopsis thaliana	
<400> 32	
Met Met Asn Pro Ser His Gly Arg Gly Leu Gly Ser Ala Gly Gly Ser	
1 5 10 15	
Ser Ser Gly Arg Asn Gln Gly Gly Gly Gly Glu Thr Val Val Glu Met	
20 25 30	
Phe Pro Ser Gly Leu Arg Val Leu Val Val Asp Asp Asp Pro Thr Cys	
35 40 45	
Leu Met Ile Leu Glu Arg Met Leu Arg Thr Cys Leu Tyr Glu Val Thr	
50 55 60	
Lys Cys Asn Arg Ala Glu Met Ala Leu Ser Leu Leu Arg Lys Asn Lys	
65 70 75 80	

## MBI0018 Sequence Listing.ST25

His Gly Phe Asp Ile Val Ile Ser Asp Val His Met Pro Asp Met Asp  
 85 90 95  
 Gly Phe Lys Leu Leu Glu His Val Gly Leu Glu Met Asp Leu Pro Val  
 100 105 110  
 Ile Met Met Ser Ala Asp Asp Ser Lys Ser Val Val Leu Lys Gly Val  
 115 120 125  
 Thr His Gly Ala Val Asp Tyr Leu Ile Lys Pro Val Arg Met Glu Ala  
 130 135 140  
 Leu Lys Asn Ile Trp Gln His Val Val Arg Lys Arg Arg Ser Glu Trp  
 145 150 155 160  
 Ser Val Pro Glu His Ser Gly Ser Ile Glu Glu Thr Gly Glu Arg Gln  
 165 170 175  
 Gln Gln Gln His Arg Gly Gly Gly Gly Gly Ala Ala Val Ser Gly Gly  
 180 185 190  
 Glu Asp Ala Val Asp Asp Asn Ser Ser Ser Val Asn Glu Gly Asn Asn  
 195 200 205  
 Trp Arg Ser Ser Ser Arg Lys Arg Lys Asp Glu Glu Gly Glu Glu Gln  
 210 215 220  
 Gly Asp Asp Lys Asp Glu Asp Ala Ser Asn Leu Lys Lys Pro Arg Val  
 225 230 235 240  
 Val Trp Ser Val Glu Leu His Gln Gln Phe Val Ala Ala Val Asn Gln  
 245 250 255  
 Leu Gly Val Glu Lys Ala Val Pro Lys Lys Ile Leu Glu Leu Met Asn  
 260 265 270  
 Val Pro Gly Leu Thr Arg Glu Asn Val Ala Ser His Leu Gln Lys Tyr  
 275 280 285  
 Arg Ile Tyr Leu Arg Arg Leu Gly Gly Val Ser Gln His Gln Gly Asn  
 290 295 300  
 Leu Asn Asn Ser Phe Met Thr Gly Gln Asp Ala Ser Phe Gly Pro Leu  
 305 310 315 320  
 Ser Thr Leu Asn Gly Phe Asp Leu Gln Ala Leu Ala Val Thr Gly Gln  
 325 330 335  
 Leu Pro Ala Gln Ser Leu Ala Gln Leu Gln Ala Ala Gly Leu Gly Arg  
 340 345 350  
 Pro Ala Met Val Ser Lys Ser Gly Leu Pro Val Ser Ser Ile Val Asp  
 355 360 365  
 Glu Arg Ser Ile Phe Ser Phe Asp Asn Thr Lys Thr Arg Phe Gly Glu  
 370 375 380

## MBI0018 Sequence Listing.ST25

Gly Leu Gly His His Gly Gln Gln Pro Gln Gln Gln Pro Gln Met Asn  
 385 390 395 400  
 Leu Leu His Gly Val Pro Thr Gly Leu Gln Gln Gln Leu Pro Met Gly  
 405 410 415  
 Asn Arg Met Ser Ile Gln Gln Gln Ile Ala Ala Val Arg Ala Gly Asn  
 420 425 430  
 Ser Val Gln Asn Asn Gly Met Leu Met Pro Leu Ala Gly Gln Gln Ser  
 435 440 445  
 Leu Pro Arg Gly Pro Pro Pro Met Leu Thr Ser Ser Gln Ser Ser Ile  
 450 455 460  
 Arg Gln Pro Met Leu Ser Asn Arg Ile Ser Glu Arg Ser Gly Phe Ser  
 465 470 475 480  
 Gly Arg Asn Asn Ile Pro Glu Ser Ser Arg Val Leu Pro Thr Ser Tyr  
 485 490 495  
 Thr Asn Leu Thr Thr Gln His Ser Ser Ser Ser Met Pro Tyr Asn Asn  
 500 505 510  
 Phe Gln Pro Glu Leu Pro Val Asn Ser Phe Pro Leu Ala Ser Ala Pro  
 515 520 525  
 Gly Ile Ser Val Pro Val Arg Lys Ala Thr Ser Tyr Gln Glu Glu Val  
 530 535 540  
 Asn Ser Ser Glu Ala Gly Phe Ile Thr Pro Ser Tyr Asp Met Phe Thr  
 545 550 555 560  
 Thr Arg Gln Asn Asp Trp Asp Leu Arg Asn Ile Gly Ile Ala Phe Asp  
 565 570 575  
 Ser His Gln Asp Ser Glu Ser Ala Ala Phe Ser Ala Ser Glu Ala Tyr  
 580 585 590  
 Ser Ser Ser Ser Met Ser Arg His Asn Thr Thr Val Ala Ala Thr Glu  
 595 600 605  
 His Gly Arg Asn His Gln Gln Pro Pro Ser Gly Met Val Gln His His  
 610 615 620  
 Gln Val Tyr Ala Asp Gly Asn Gly Gly Ser Val Arg Val Lys Ser Glu  
 625 630 635 640  
 Arg Val Ala Thr Asp Thr Ala Thr Met Ala Phe His Glu Gln Tyr Ser  
 645 650 655  
 Asn Gln Glu Asp Leu Met Ser Ala Leu Leu Lys Gln Val  
 660 665



## MBI0018 Sequence Listing.ST25

```

<211> 1239
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (6)..(1091)
<223> G993

<400> 33
caaat atg gaa tac agc tgt gta gac gac agt agt aca acg tca gaa tct 50
      Met Glu Tyr Ser Cys Val Asp Asp Ser Ser Thr Thr Ser Glu Ser
      1          5          10          15

ctc tcc atc tct act act cca aag ccg aca acg acg acg gag aag aaa 98
Leu Ser Ile Ser Thr Thr Pro Lys Pro Thr Thr Thr Thr Glu Lys Lys
      20          25          30

ctc tct tct ccg ccg gcg acg tcg atg cgt ctc tac aga atg gga agc 146
Leu Ser Ser Pro Pro Ala Thr Ser Met Arg Leu Tyr Arg Met Gly Ser
      35          40          45

ggc gga agc agc gtc gtt ttg gat tca gag aac ggc gtc gag acc gag 194
Gly Gly Ser Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Thr Glu
      50          55          60

tca cgt aag ctt cct tcg tcg aaa tat aaa ggc gtt gtg cct cag cct 242
Ser Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly Val Val Pro Gln Pro
      65          70          75

aac gga aga tgg gga gct cag att tac gag aag cat cag cga gtt tgg 290
Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys His Gln Arg Val Trp
      80          85          90          95

ctc ggt act ttc aac gag gaa gaa gaa gct gcg tct tct tac gac atc 338
Leu Gly Thr Phe Asn Glu Glu Glu Glu Ala Ala Ser Ser Tyr Asp Ile
      100          105          110

gcc gtg agg aga ttc cgc ggc cgc gac gcc gtc act aac ttc aaa tct 386
Ala Val Arg Arg Phe Arg Gly Arg Asp Ala Val Thr Asn Phe Lys Ser
      115          120

caa gtt gat gga aac gac gcc gaa tcg gct ttt ctt gac gct cat tct 434
Gln Val Asp Gly Asn Asp Ala Glu Ser Ala Phe Leu Asp Ala His Ser
      130          135          140

aaa gct gag atc gtg gat atg ttg agg aaa cac act tac gcc gat gag 482
Lys Ala Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr Ala Asp Glu
      145          150          155

ttt gag cag agt aga cgg aag ttt gtt aac ggc gac gga aaa cgc tct 530
Phe Glu Gln Ser Arg Arg Lys Phe Val Asn Gly Asp Gly Lys Arg Ser
      160          165          170          175

ggg ttg gag acg gcg acg tac gga aac gac gct gtt ttg aga gcg cgt 578
Gly Leu Glu Thr Ala Thr Tyr Gly Asn Asp Ala Val Leu Arg Ala Arg
      180          185          190

gag gtt ttg ttc gag aag act gtt acg ccg agc gac gtc ggg aag ctg 626
Glu Val Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val Gly Lys Leu
      195          200          205

aac cgt tta gtg ata ccg aaa caa cac gcg gag aag cat ttt ccg tta 674
Asn Arg Leu Val Ile Pro Lys Gln His Ala Glu Lys His Phe Pro Leu
      210          215          220

ccg gcg atg acg acg gcg atg ggg atg aat ccg tct ccg acg aaa ggc 722
Pro Ala Met Thr Thr Ala Met Gly Met Asn Pro Ser Pro Thr Lys Gly
      225          230          235

gtt ttg att aac ttg gaa gat aga aca ggg aaa gtg tgg cgg ttc cgt 770
Val Leu Ile Asn Leu Glu Asp Arg Thr Gly Lys Val Trp Arg Phe Arg
      240          245          250          255

```

MBI0018 Sequence Listing.ST25

tac agt tac tgg aac agc agt caa agt tac gtg ttg acc aag ggc tgg	818
Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp	
260 265 270	
agc cgg ttc gtt aaa gag aag aat ctt cga gcc ggt gat gtg gtt tgt	866
Ser Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val Cys	
275 280 285	
ttc gag aga tca acc gga cca gac cgg caa ttg tat atc cac tgg aaa	914
Phe Glu Arg Ser Thr Gly Pro Asp Arg Gln Leu Tyr Ile His Trp Lys	
290 295 300	
gtc cgg tct agt ccg gtt cag act gtg gtt agg cta ttc gga gtc aac	962
Val Arg Ser Ser Pro Val Gln Thr Val Val Arg Leu Phe Gly Val Asn	
305 310 315	
att ttc aat gtg agt aac gag aaa cca aac gac gtc gca gta gag tgt	1010
Ile Phe Asn Val Ser Asn Glu Lys Pro Asn Asp Val Ala Val Glu Cys	
320 325 330 335	
gtt ggc aag aag aga tct cgg gaa gat gat ttg ttt tcg tta ggg tgt	1058
Val Gly Lys Lys Arg Ser Arg Glu Asp Asp Leu Phe Ser Leu Gly Cys	
340 345 350	
tcc aag aag cag gcg att atc aac atc ttg tga caaattcttt ttttttggtt	1111
Ser Lys Lys Gln Ala Ile Ile Asn Ile Leu	
355 360	
tttttcttca atttgtttct cttttttcaa tattttgtat tgaaatgaca agttgtaaat	1171
taggacaaga caagaaaaaa tgacaactag acaaaatagt ttttgtttaa aaaaaaaaaa	1231
aaaaaaaaa	1239
<210> 34	
<211> 361	
<212> PRT	
<213> Arabidopsis thaliana	
<400> 34	
Met Glu Tyr Ser Cys Val Asp Asp Ser Ser Thr Thr Ser Glu Ser Leu	
1 5 10 15	
Ser Ile Ser Thr Thr Pro Lys Pro Thr Thr Thr Thr Glu Lys Lys Leu	
20 25 30	
Ser Ser Pro Pro Ala Thr Ser Met Arg Leu Tyr Arg Met Gly Ser Gly	
35 40 45	
Gly Ser Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Thr Glu Ser	
50 55 60	
Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly Val Val Pro Gln Pro Asn	
65 70 75 80	
Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys His Gln Arg Val Trp Leu	
85 90 95	
Gly Thr Phe Asn Glu Glu Glu Glu Ala Ala Ser Ser Tyr Asp Ile Ala	
100 105 110	
Val Arg Arg Phe Arg Gly Arg Asp Ala Val Thr Asn Phe Lys Ser Gln	
115 120 125	
Val Asp Gly Asn Asp Ala Glu Ser Ala Phe Leu Asp Ala His Ser Lys	

MBI0018 Sequence Listing.ST25

130	135	140
Ala Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr Ala Asp Glu Phe		
145	150	155 160
Glu Gln Ser Arg Arg Lys Phe Val Asn Gly Asp Gly Lys Arg Ser Gly		
	165	170 175
Leu Glu Thr Ala Thr Tyr Gly Asn Asp Ala Val Leu Arg Ala Arg Glu		
	180	185 190
Val Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val Gly Lys Leu Asn		
	195	200 205
Arg Leu Val Ile Pro Lys Gln His Ala Glu Lys His Phe Pro Leu Pro		
	210	215 220
Ala Met Thr Thr Ala Met Gly Met Asn Pro Ser Pro Thr Lys Gly Val		
	225	230 235 240
Leu Ile Asn Leu Glu Asp Arg Thr Gly Lys Val Trp Arg Phe Arg Tyr		
	245	250 255
Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp Ser		
	260	265 270
Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val Cys Phe		
	275	280 285
Glu Arg Ser Thr Gly Pro Asp Arg Gln Leu Tyr Ile His Trp Lys Val		
	290	295 300
Arg Ser Ser Pro Val Gln Thr Val Val Arg Leu Phe Gly Val Asn Ile		
	305	310 315 320
Phe Asn Val Ser Asn Glu Lys Pro Asn Asp Val Ala Val Glu Cys Val		
	325	330 335
Gly Lys Lys Arg Ser Arg Glu Asp Asp Leu Phe Ser Leu Gly Cys Ser		
	340	345 350
Lys Lys Gln Ala Ile Ile Asn Ile Leu		
	355	360

<210> 35  
 <211> 1281  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (64)..(1098)  
 <223> G867

<400> 35	
cacaacacaa acacatttct gttttctcca ttgtttcaaa ccataaaaaa aaacacagat	60
taa atg gaa tcg agt agc gtt gat gag agt act aca agt aca ggt tcc	108
Met Glu Ser Ser Ser Val Asp Glu Ser Thr Thr Ser Thr Gly Ser	

MBI0018 Sequence Listing.ST25																
1				5				10				15				
atc Ile	tgt Cys	gaa Glu	acc Thr	ccg Pro 20	gcg Ala	ata Ile	act Thr	ccg Pro	gcg Ala 25	aaa Lys	aag Lys	tcg Ser	tcg Ser	gta Val 30	ggt Gly	156
aac Asn	tta Leu	tac Tyr	agg Arg 35	atg Met	gga Gly	agc Ser	gga Gly	tca Ser 40	agc Ser	gtt Val	gtg Val	tta Leu	gat Asp 45	tca Ser	gag Glu	204
aac Asn	ggc Gly	gta Val 50	gaa Glu	gct Ala	gaa Glu	tct Ser	agg Arg 55	aag Lys	ctt Leu	ccg Pro	tcg Ser	tca Ser 60	aaa Lys	tac Tyr	aaa Lys	252
ggt Gly	gtg Val 65	gtg Val	cca Pro	caa Gln	cca Pro	aac Asn 70	gga Gly	aga Arg	tgg Trp	gga Gly	gct Ala 75	cag Gln	att Ile	tac Tyr	gag Glu	300
aaa Lys 80	cac His	cag Gln	cgc Arg	gtg Val	tgg Trp 85	ctc Leu	ggg Gly	aca Thr	ttc Phe	aac Asn 90	gaa Glu	gaa Glu	gac Asp	gaa Glu	gcc Ala 95	348
gct Ala	cgt Arg	gcc Ala	tac Tyr	gac Asp 100	gtc Val	gcg Ala	gtt Val	cac His	agg Arg 105	ttc Phe	cgt Arg	cgc Arg	cgt Arg	gac Asp 110	gcc Ala	396
gtc Val	aca Thr	aat Asn	ttc Phe 115	aaa Lys	gac Asp	gtg Val	aag Lys	atg Met 120	gac Asp	gaa Glu	gac Asp	gag Glu	gtc Val 125	gat Asp	ttc Phe	444
ttg Leu	aat Asn	tct Ser 130	cat His	tcg Ser	aaa Lys	tct Ser	gag Glu 135	atc Ile	gtt Val	gat Asp	atg Met	ttg Leu 140	agg Arg	aaa Lys	cat His	492
act Thr	tat Tyr 145	aac Asn	gaa Glu	gag Glu	tta Leu	gag Glu 150	cag Gln	agt Ser	aaa Lys	cgg Arg	cgt Arg 155	cgt Arg	aat Asn	ggt Gly	aac Asn	540
gga Gly 160	aac Asn	atg Met	act Thr	agg Arg	acg Thr 165	ttg Leu	tta Leu	acg Thr	tcg Ser	ggg Gly 170	ttg Leu	agt Ser	aat Asn	gat Asp	ggt Gly 175	588
gtt Val	tct Ser	acg Thr	acg Thr	ggg Gly 180	ttt Phe	aga Arg	tcg Ser	gcg Ala	gag Glu 185	gca Ala	ctg Leu	ttt Phe	gag Glu	aaa Lys 190	gcg Ala	636
gta Val	acg Thr	cca Pro	agc Ser 195	gac Asp	gtt Val	ggg Gly	aag Lys	cta Leu 200	aac Asn	cgt Arg	ttg Leu	gtt Val	ata Ile 205	ccg Pro	aaa Lys	684
cat His	cac His	gca Ala 210	gag Glu	aaa Lys	cat His	ttt Phe	ccg Pro 215	tta Leu	ccg Pro	tca Ser	agt Ser	aac Asn 220	gtt Val	tcc Ser	gtg Val	732
aaa Lys	gga Gly 225	gtg Val	ttg Leu	ttg Leu	aac Asn	ttt Phe 230	gag Glu	gac Asp	gtt Val	aac Asn	ggg Gly 235	aaa Lys	gtg Val	tgg Trp	agg Arg	780
ttc Phe 240	cgt Arg	tac Tyr	tcg Ser	tat Tyr	tgg Trp 245	aac Asn	agt Ser	agt Ser	cag Gln 250	agt Ser	tat Tyr	gtt Val	ttg Leu	act Thr 255	aaa Lys	828
ggt Gly	tgg Trp	agc Ser	agg Arg	ttc Phe 260	gtt Val	aag Lys	gag Glu	aag Lys	aat Asn 265	cta Leu	cgt Arg	gct Ala	ggt Gly	gac Asp 270	gtg Val	876
gtt Val	agt Ser	ttc Phe	agt Ser 275	aga Arg	tct Ser	aac Asn	ggt Gly	cag Gln 280	gat Asp	caa Gln	cag Gln	ttg Leu	tac Tyr 285	att Ile	ggg Gly	924
tgg Trp	aag Lys	tcg Ser 290	aga Arg	tcc Ser	ggg Gly	tca Ser	gat Asp 295	tta Leu	gat Asp	gcg Ala	ggt Gly	cgg Arg 300	gtt Val	ttg Leu	aga Arg	972
ttg	ttc	gga	gtt	aac	att	tca	ccg	gag	agt	tca	aga	aac	gac	gtc	gta	1020

MBI0018 Sequence Listing.ST25

```

Leu Phe Gly Val Asn Ile Ser Pro Glu Ser Ser Arg Asn Asp Val Val
 305                      310                      315

gga aac aaa aga gtg aac gat act gag atg tta tcg ttg gtg tgt agc      1068
Gly Asn Lys Arg Val Asn Asp Thr Glu Met Leu Ser Leu Val Cys Ser
320                      325                      330                      335

aag aag caa cgc atc ttt cac gcc tcg taa caactcttct tctttttttt      1118
Lys Lys Gln Arg Ile Phe His Ala Ser
                      340

tcttttggtg ttttaataat ttttaaaaac tccatttttcg ttttctttat ttgcatcggt      1178

ttctttcttc ttgtttacca aaggttcatg agttgttttt gttgtattga tgaactgtaa      1238

attttattta taggataaat ttttaaaaaa aaaaaaaaaa aaa                        1281

<210> 36
<211> 344
<212> PRT
<213> Arabidopsis thaliana

<400> 36

Met Glu Ser Ser Ser Val Asp Glu Ser Thr Thr Ser Thr Gly Ser Ile
1                      5                      10                      15

Cys Glu Thr Pro Ala Ile Thr Pro Ala Lys Lys Ser Ser Val Gly Asn
20                      25                      30

Leu Tyr Arg Met Gly Ser Gly Ser Ser Val Val Leu Asp Ser Glu Asn
35                      40                      45

Gly Val Glu Ala Glu Ser Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly
50                      55                      60

Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys
65                      70                      75                      80

His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Glu Asp Glu Ala Ala
85                      90                      95

Arg Ala Tyr Asp Val Ala Val His Arg Phe Arg Arg Arg Asp Ala Val
100                      105                      110

Thr Asn Phe Lys Asp Val Lys Met Asp Glu Asp Glu Val Asp Phe Leu
115                      120                      125

Asn Ser His Ser Lys Ser Glu Ile Val Asp Met Leu Arg Lys His Thr
130                      135                      140

Tyr Asn Glu Glu Leu Glu Gln Ser Lys Arg Arg Arg Asn Gly Asn Gly
145                      150                      155                      160

Asn Met Thr Arg Thr Leu Leu Thr Ser Gly Leu Ser Asn Asp Gly Val
165                      170                      175

Ser Thr Thr Gly Phe Arg Ser Ala Glu Ala Leu Phe Glu Lys Ala Val
180                      185                      190

Thr Pro Ser Asp Val Gly Lys Leu Asn Arg Leu Val Ile Pro Lys His
195                      200                      205

```

## MBI0018 Sequence Listing.ST25

His Ala Glu Lys His Phe Pro Leu Pro Ser Ser Asn Val Ser Val Lys  
 210 215 220

Gly Val Leu Leu Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg Phe  
 225 230 235 240

Arg Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly  
 245 250 255

Trp Ser Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val  
 260 265 270

Ser Phe Ser Arg Ser Asn Gly Gln Asp Gln Gln Leu Tyr Ile Gly Trp  
 275 280 285

Lys Ser Arg Ser Gly Ser Asp Leu Asp Ala Gly Arg Val Leu Arg Leu  
 290 295 300

Phe Gly Val Asn Ile Ser Pro Glu Ser Ser Arg Asn Asp Val Val Gly  
 305 310 315 320

Asn Lys Arg Val Asn Asp Thr Glu Met Leu Ser Leu Val Cys Ser Lys  
 325 330 335

Lys Gln Arg Ile Phe His Ala Ser  
 340

<210> 37  
 <211> 1155  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (76)..(1077)  
 <223> G1930

<400> 37  
 attcacatta ctaatctctc aagatttcac aattttcttg tgattttctc tcagtttctt 60

atttcgtttc ataac atg gat gcc atg agt agc gta gac gag agc tct aca 111  
 Met Asp Ala Met Ser Ser Val Asp Glu Ser Ser Thr  
 1 5 10

act aca gat tcc att ccg gcg aga aag tca tcg tct ccg gcg agt tta 159  
 Thr Thr Asp Ser Ile Pro Ala Arg Lys Ser Ser Ser Pro Ala Ser Leu  
 15 20 25

cta tat aga atg gga agc gga aca agc gtg gta ctt gat tca gag aac 207  
 Leu Tyr Arg Met Gly Ser Gly Thr Ser Val Val Leu Asp Ser Glu Asn  
 30 35 40

ggt gtc gaa gtc gaa gtc gaa gcc gaa tca aga aag ctt cct tct tca 255  
 Gly Val Glu Val Glu Val Glu Ala Glu Ser Arg Lys Leu Pro Ser Ser  
 45 50 55 60

aga ttc aaa ggt gtt gtt cct caa cca aat gga aga tgg gga gct cag 303  
 Arg Phe Lys Gly Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln  
 65 70 75

att tac gag aaa cat caa cgc gtg tgg ctt ggt act ttc aac gag gaa 351  
 Ile Tyr Glu Lys His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Glu  
 80 85 90

## MBI0018 Sequence Listing.ST25

gac gaa gca gct cgt gct tac gac gtc gcg gct cac cgt ttc cgt ggc Asp Glu Ala Ala Arg Ala Tyr Asp Val Ala Ala His Arg Phe Arg Gly 95 100 105	399
cgc gat gcc gtt act aat ttc aaa gac acg acg ttc gaa gaa gag gtt Arg Asp Ala Val Thr Asn Phe Lys Asp Thr Thr Phe Glu Glu Glu Val 110 115 120	447
gag ttc tta aac gcg cat tcg aaa tca gag atc gta gat atg ttg aga Glu Phe Leu Asn Ala His Ser Lys Ser Glu Ile Val Asp Met Leu Arg 125 130 135 140	495
aaa cac act tac aaa gaa gag tta gac caa agg aaa cgt aac cgt gac Lys His Thr Tyr Lys Glu Glu Leu Asp Gln Arg Lys Arg Asn Arg Asp 145 150 155	543
ggg aac gga aaa gag acg acg gcg ttt gct ttg gct tcg atg gtg gtt Gly Asn Gly Lys Glu Thr Thr Ala Phe Ala Leu Ala Ser Met Val Val 160 165 170	591
atg acg ggg ttt aaa acg gcg gag tta ctg ttt gag aaa acg gta acg Met Thr Gly Phe Lys Thr Ala Glu Leu Leu Phe Glu Lys Thr Val Thr 175 180 185	639
cca agt gac gtc ggg aaa cta aac cgt tta gtt ata cca aaa cac caa Pro Ser Asp Val Gly Lys Leu Asn Arg Leu Val Ile Pro Lys His Gln 190 195 200	687
gcg gag aaa cat ttt ccg tta ccg tta ggt aat aat aac gtc tcc gtt Ala Glu Lys His Phe Pro Leu Pro Leu Gly Asn Asn Asn Val Ser Val 205 210 220	735
aaa ggt atg ctg ttg aat ttc gaa gac gtt aac ggg aaa gtg tgg agg Lys Gly Met Leu Leu Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg 225 230 235	783
ttc cgt tac tct tat tgg aat agt agt caa agt tat gtg ttg acc aaa Phe Arg Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys 240 245 250	831
ggg tgg agt aga ttc gtt aaa gag aag aga ctt tgt gct ggt gat ttg Gly Trp Ser Arg Phe Val Lys Glu Lys Arg Leu Cys Ala Gly Asp Leu 255 260 265	879
atc agt ttt aaa aga tcc aac gat caa gat caa aaa ttc ttt atc ggg Ile Ser Phe Lys Arg Ser Asn Asp Gln Asp Gln Lys Phe Phe Ile Gly 270 275 280	927
tgg aaa tcg aaa tcc ggg ttg gat cta gag acg ggt cgg gtt atg aga Trp Lys Ser Lys Ser Gly Leu Asp Leu Glu Thr Gly Arg Val Met Arg 285 290 295 300	975
ttg ttt ggg gtt gat att tct tta aac gcc gtc gtt gta gtg aag gaa Leu Phe Gly Val Asp Ile Ser Leu Asn Ala Val Val Val Val Lys Glu 305 310 315	1023
aca acg gag gtg tta atg tcg tcg tta agg tgt aag aag caa cga gtt Thr Thr Glu Val Leu Met Ser Ser Leu Arg Cys Lys Lys Gln Arg Val 320 325 330	1071
ttg taa taacaattta acaacttggg aaagaaaaaa aagctttttg attttaattt Leu	1127
ctcttcaacg ttaatcttgc tgagatta	1155
<210> 38	
<211> 333	
<212> PRT	
<213> Arabidopsis thaliana	
<400> 38	

## MBI0018 Sequence Listing.ST25

Met Asp Ala Met Ser Ser Val Asp Glu Ser Ser Thr Thr Thr Asp Ser  
 1 5 10 15  
 Ile Pro Ala Arg Lys Ser Ser Ser Pro Ala Ser Leu Leu Tyr Arg Met  
 20 25 30  
 Gly Ser Gly Thr Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Val  
 35 40 45  
 Glu Val Glu Ala Glu Ser Arg Lys Leu Pro Ser Ser Arg Phe Lys Gly  
 50 55 60  
 Val Val Pro Gln Pro Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys  
 65 70 75 80  
 His Gln Arg Val Trp Leu Gly Thr Phe Asn Glu Glu Asp Glu Ala Ala  
 85 90 95  
 Arg Ala Tyr Asp Val Ala Ala His Arg Phe Arg Gly Arg Asp Ala Val  
 100 105 110  
 Thr Asn Phe Lys Asp Thr Thr Phe Glu Glu Glu Val Glu Phe Leu Asn  
 115 120 125  
 Ala His Ser Lys Ser Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr  
 130 135 140  
 Lys Glu Glu Leu Asp Gln Arg Lys Arg Asn Arg Asp Gly Asn Gly Lys  
 145 150 155 160  
 Glu Thr Thr Ala Phe Ala Leu Ala Ser Met Val Val Met Thr Gly Phe  
 165 170 175  
 Lys Thr Ala Glu Leu Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val  
 180 185 190  
 Gly Lys Leu Asn Arg Leu Val Ile Pro Lys His Gln Ala Glu Lys His  
 195 200 205  
 Phe Pro Leu Pro Leu Gly Asn Asn Asn Val Ser Val Lys Gly Met Leu  
 210 215 220  
 Leu Asn Phe Glu Asp Val Asn Gly Lys Val Trp Arg Phe Arg Tyr Ser  
 225 230 235 240  
 Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp Ser Arg  
 245 250 255  
 Phe Val Lys Glu Lys Arg Leu Cys Ala Gly Asp Leu Ile Ser Phe Lys  
 260 265 270  
 Arg Ser Asn Asp Gln Asp Gln Lys Phe Phe Ile Gly Trp Lys Ser Lys  
 275 280 285  
 Ser Gly Leu Asp Leu Glu Thr Gly Arg Val Met Arg Leu Phe Gly Val  
 290 295 300



## MBI0018 Sequence Listing.ST25

Asp Ile Ser Leu Asn Ala Val Val Val Val Lys Glu Thr Thr Glu Val  
305 310 315 320

Leu Met Ser Ser Leu Arg Cys Lys Lys Gln Arg Val Leu  
325 330

<210> 39  
<211> 1035  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (1)..(1035)  
<223> G1594

<400> 39  
atg tac aat ttc cat tcg gcc ggt gat tat tca gat aag tcg gtt ctg 48  
Met Tyr Asn Phe His Ser Ala Gly Asp Tyr Ser Asp Lys Ser Val Leu  
1 5 10 15  
atg atg tca ccg gag agt ctc atg ttt cct tcc gat tac caa gct ttg 96  
Met Met Ser Pro Glu Ser Leu Met Phe Pro Ser Asp Tyr Gln Ala Leu  
20 25 30  
cta tgt tcc tcc gcc ggt gaa aat cgt gtc tct gat gtt ttc gga tcc 144  
Leu Cys Ser Ser Ala Gly Glu Asn Arg Val Ser Asp Val Phe Gly Ser  
35 40 45  
gac gag cta ctc tca gta gcc gtc tcc gct ttg tcg tcg gag gcg gct 192  
Asp Glu Leu Leu Ser Val Ala Val Ser Ala Leu Ser Ser Glu Ala Ala  
50 55 60  
tcg atc gct ccg gag atc cga aga aat gat gat aac gtt tct cta act 240  
Ser Ile Ala Pro Glu Ile Arg Arg Asn Asp Asp Asn Val Ser Leu Thr  
65 70 75 80  
gtc atc aaa gct aaa atc gct tgt cat cct tcg tat cct cgc tta ctt 288  
Val Ile Lys Ala Lys Ile Ala Cys His Pro Ser Tyr Pro Arg Leu Leu  
85 90 95  
caa gct tac atc gat tgc caa aag gtc gga gca cca ccg gag ata gcg 336  
Gln Ala Tyr Ile Asp Cys Gln Lys Val Gly Ala Pro Pro Glu Ile Ala  
100 105 110  
tgt tta cta gag gag att caa cgg gag agt gat gtt tat aag caa gag 384  
Cys Leu Leu Glu Glu Ile Gln Arg Glu Ser Asp Val Tyr Lys Gln Glu  
115 120 125  
gtt gtt cct tct tct tgc ttt gga gct gat cct gag ctt gat gaa ttt 432  
Val Val Pro Ser Ser Cys Phe Gly Ala Asp Pro Glu Leu Asp Glu Phe  
130 135 140  
atg gaa acg tac tgc gat ata tta gtg aaa tac aaa tcg gat cta gca 480  
Met Glu Thr Tyr Cys Asp Ile Leu Val Lys Tyr Lys Ser Asp Leu Ala  
145 150 155 160  
aga ccg ttt gac gag gca acg tgt ttc ttg aac aag att gag atg cag 528  
Arg Pro Phe Asp Glu Ala Thr Cys Phe Leu Asn Lys Ile Glu Met Gln  
165 170 175  
cta cgg aac cta tgt act ggt gtc gag tct gcc agg gga gtt tct ggg 576  
Leu Arg Asn Leu Cys Thr Gly Val Glu Ser Ala Arg Gly Val Ser Gly  
180 185 190  
ggg atg tct cct cat ggg gac aag act att agt cct ctc ctg aca aat 624  
Gly Met Ser Pro His Gly Asp Lys Thr Ile Ser Pro Leu Thr Asn  
195 200 205  
gac aat gga gag gat ggt gta ata tca tct gac gag gaa ctg agt gga 672  
Asp Asn Gly Glu Asp Gly Val Ile Ser Ser Asp Glu Glu Leu Ser Gly

## MBI0018 Sequence Listing.ST25

210	215	220	
ggt gat cat gag gta gca gag gat ggg aga caa aga tgt gaa gac cgg			720
Gly Asp His Glu Val Ala Glu Asp Gly Arg Gln Arg Cys Glu Asp Arg	230	235	240
225			
gac ctc aaa gat agg ttg cta cgc aaa ttt gga agc cgt att agt act			768
Asp Leu Lys Asp Arg Leu Leu Arg Lys Phe Gly Ser Arg Ile Ser Thr	245	250	255
260			
tta aag ctt gag ttc tca aag aag aag aag aaa gga aag tta cca aga			816
Leu Lys Leu Glu Phe Ser Lys Lys Lys Lys Lys Gly Lys Leu Pro Arg	265	270	
275			
gaa gca aga caa gct ctt ctt gat tgg tgg aat ctc cat tat aag tgg			864
Glu Ala Arg Gln Ala Leu Leu Asp Trp Trp Asn Leu His Tyr Lys Trp	280	285	
290			
cct tac cct act gaa gga gat aag ata gca tta gct gat gca acg ggg			912
Pro Tyr Pro Thr Glu Gly Asp Lys Ile Ala Leu Ala Asp Ala Thr Gly	295	300	
305			
tta gac caa aaa caa atc aac aat tgg ttt ata aac caa agg aaa cgt			960
Leu Asp Gln Lys Gln Ile Asn Asn Trp Phe Ile Asn Gln Arg Lys Arg	310	315	320
325			
cat tgg aag cca tca gag aat atg cct ttc gct atg atg gat gat tct			1008
His Trp Lys Pro Ser Glu Asn Met Pro Phe Ala Met Met Asp Asp Ser	330	335	
340			
agt gga tca ttc ttt acc gag gaa tga			1035
Ser Gly Ser Phe Phe Thr Glu Glu			

<210> 40  
 <211> 344  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 40

Met Tyr Asn Phe His Ser Ala Gly Asp Tyr Ser Asp Lys Ser Val Leu	1	5	10	15
Met Met Ser Pro Glu Ser Leu Met Phe Pro Ser Asp Tyr Gln Ala Leu	20	25	30	
Leu Cys Ser Ser Ala Gly Glu Asn Arg Val Ser Asp Val Phe Gly Ser	35	40	45	
Asp Glu Leu Leu Ser Val Ala Val Ser Ala Leu Ser Ser Glu Ala Ala	50	55	60	
Ser Ile Ala Pro Glu Ile Arg Arg Asn Asp Asp Asn Val Ser Leu Thr	65	70	75	80
Val Ile Lys Ala Lys Ile Ala Cys His Pro Ser Tyr Pro Arg Leu Leu	85	90	95	
Gln Ala Tyr Ile Asp Cys Gln Lys Val Gly Ala Pro Pro Glu Ile Ala	100	105	110	
Cys Leu Leu Glu Glu Ile Gln Arg Glu Ser Asp Val Tyr Lys Gln Glu	115	120	125	
Val Val Pro Ser Ser Cys Phe Gly Ala Asp Pro Glu Leu Asp Glu Phe				

MBI0018 Sequence Listing.ST25

130	135	140	
Met Glu Thr Tyr Cys Asp Ile Leu Val Lys Tyr Lys Ser Asp Leu Ala			
145	150	155	160
Arg Pro Phe Asp Glu Ala Thr Cys Phe Leu Asn Lys Ile Glu Met Gln			
	165	170	175
Leu Arg Asn Leu Cys Thr Gly Val Glu Ser Ala Arg Gly Val Ser Gly			
	180	185	190
Gly Met Ser Pro His Gly Asp Lys Thr Ile Ser Pro Leu Leu Thr Asn			
	195	200	205
Asp Asn Gly Glu Asp Gly Val Ile Ser Ser Asp Glu Glu Leu Ser Gly			
	210	215	220
Gly Asp His Glu Val Ala Glu Asp Gly Arg Gln Arg Cys Glu Asp Arg			
	225	230	235
Asp Leu Lys Asp Arg Leu Leu Arg Lys Phe Gly Ser Arg Ile Ser Thr			
	245	250	255
Leu Lys Leu Glu Phe Ser Lys Lys Lys Lys Gly Lys Leu Pro Arg			
	260	265	270
Glu Ala Arg Gln Ala Leu Leu Asp Trp Trp Asn Leu His Tyr Lys Trp			
	275	280	285
Pro Tyr Pro Thr Glu Gly Asp Lys Ile Ala Leu Ala Asp Ala Thr Gly			
	290	295	300
Leu Asp Gln Lys Gln Ile Asn Asn Trp Phe Ile Asn Gln Arg Lys Arg			
	305	310	315
His Trp Lys Pro Ser Glu Asn Met Pro Phe Ala Met Met Asp Asp Ser			
	325	330	335
Ser Gly Ser Phe Phe Thr Glu Glu			
	340		

<210> 41  
 <211> 2559  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)..(2259)  
 <223> G391

<400> 41			
atg atg atg gtc cat tcg atg agc aga gat atg atg aac aga gag tcg			48
Met Met Met Val His Ser Met Ser Arg Asp Met Met Asn Arg Glu Ser			
1	5	10	15
ccg gat aaa ggg tta gat tcc ggc aag tat gtg agg tac acg ccg gag			96
Pro Asp Lys Gly Leu Asp Ser Gly Lys Tyr Val Arg Tyr Thr Pro Glu			
	20	25	30

MBI0018 Sequence Listing.ST25																		
caa Gln	gtg Val	gaa Glu 35	gct Ala	ctc Leu	gag Glu	aga Arg	gtt Val 40	tac Tyr	act Thr	gag Glu	tgt Cys	cct Pro 45	aag Lys	cca Pro	agt Ser	144		
tct Ser	cta Leu 50	aga Arg	aga Arg	caa Gln	caa Gln	ctc Leu 55	ata Ile	cgt Arg	gaa Glu	tgt Cys	ccg Pro 60	att Ile	ctc Leu	tct Ser	aac Asn	192		
atc Ile 65	gag Glu	cct Pro	aag Lys	cag Gln	atc Ile 70	aaa Lys	gtt Val	tgg Trp	ttt Phe	cag Gln 75	aac Asn	cgc Arg	aga Arg	tgt Cys	cgt Arg 80	240		
gag Glu	aag Lys	cag Gln	agg Arg	aaa Lys 85	gaa Glu	gct Ala	gct Ala	cgt Arg	ctt Leu 90	caa Gln	aca Thr	gtg Val	aac Asn	aga Arg 95	aaa Lys	288		
ctc Leu	aat Asn	gcc Ala	atg Met 100	aac Asn	aaa Lys	ctc Leu	ttg Leu	atg Met 105	gaa Glu	gag Glu	aat Asn	gat Asp	cgt Arg 110	ttg Leu	cag Gln	336		
aag Lys	caa Gln	gtt Val 115	tct Ser	aac Asn	ttg Leu	gtc Val	tat Tyr 120	gag Glu	aat Asn	ggc Gly	cac His	atg Met 125	aaa Lys	cat His	caa Gln	384		
ctt Leu	cac His 130	act Thr	gct Ala	tct Ser	ggg Gly	acg Thr 135	acc Thr	aca Thr	gac Asp	aac Asn	agc Ser 140	tgt Cys	gag Glu	tct Ser	gtg Val	432		
gtc Val 145	gtg Val	agt Ser	ggt Gly	cag Gln	caa Gln 150	cat His	caa Gln	cag Gln	caa Gln	aac Asn 155	cca Pro	aat Asn	cct Pro	cag Gln	cat His 160	480		
cag Gln	caa Gln	cgt Arg	gat Asp	gct Ala 165	aac Asn	aac Asn	cca Pro	gca Ala	gga Gly 170	ctc Leu	ctt Leu	tct Ser	ata Ile	gca Ala 175	gag Glu	528		
gag Glu	gcc Ala	cta Leu 180	gca Ala	gag Glu	ttc Phe	ctt Leu	tcc Ser	aag Lys 185	gct Ala	aca Thr	gga Gly	act Thr	gct Ala 190	gtt Val	gac Asp	576		
tgg Trp	gtt Val	cag Gln 195	atg Met	att Ile	ggg Gly	atg Met	aag Lys 200	cct Pro	ggt Gly	ccg Pro	gat Asp	tct Ser 205	att Ile	ggc Gly	ata Ile	624		
gtc Val	gct Ala 210	att Ile	tcg Ser	cgc Arg	aac Asn	tgc Cys 215	agc Ser	gga Gly	att Ile	gca Ala	gca Ala 220	cgt Arg	gcc Ala	tgc Cys	ggc Gly	672		
ctc Leu 225	gtg Val	agt Ser	tta Leu	gaa Glu	ccc Pro 230	atg Met	aag Lys	gtt Val	gct Ala	gaa Glu 235	att Ile	ctc Leu	aaa Lys	gat Asp	cgt Arg 240	720		
cca Pro	tct Ser	tgg Trp	ctc Leu	cga Arg 245	gat Asp	tgt Cys	cga Arg	agt Ser	gtg Val 250	gat Asp	act Thr	ctg Leu	agt Ser	gtg Val 255	ata Ile	768		
cct Pro	gct Ala	gga Gly	aac Asn 260	ggt Gly	ggg Gly	acg Thr	atc Ile	gag Glu 265	ctt Leu	att Ile	tac Tyr	acg Thr	cag Gln 270	atg Met	tat Tyr	816		
gct Ala	cct Pro	acg Thr 275	act Thr	tta Leu	gca Ala	gca Ala	gct Ala 280	cgt Arg	gac Asp	ttt Phe	tgg Trp	acg Thr 285	ctg Leu	aga Arg	tat Tyr	864		
agc Ser	aca Thr 290	tgt Cys	ttg Leu	gaa Glu	gat Asp	gga Gly 295	agc Ser	tat Tyr	gtg Val	gtt Val	tgt Cys 300	gaa Glu	agg Arg	tcg Ser	ctt Leu	912		
act Thr 305	tct Ser	gca Ala	act Thr	ggt Gly	ggc Gly 310	ccc Pro	act Thr	ggg Gly	cca Pro	cct Pro 315	tct Ser	tca Ser	aac Asn	ttt Phe	gtg Val 320	960		
aga Arg	gct Ala	gaa Glu	atg Met	aaa Lys 325	cca Pro	agc Ser	ggg Gly	ttt Phe	ctc Leu 330	atc Ile	cgt Arg	cct Pro	tgc Cys	gat Asp 335	ggc Gly	1008		

## MBI0018 Sequence Listing.ST25

ggt ggt tcc att ctc cac att gtt gat cat gtt gat ctg gat gcc tgg	1056
Gly Gly Ser Ile Leu His Ile Val Asp His Val Asp Leu Asp Ala Trp	
340 345 350	
agt gtc cct gaa gtc atg agg cct ctc tat gaa tca tcg aag att ctt	1104
Ser Val Pro Glu Val Met Arg Pro Leu Tyr Glu Ser Lys Ile Leu	
355 360 365	
gct cag aaa atg act gtt gct gct ttg aga cat gta aga caa att gca	1152
Ala Gln Lys Met Thr Val Ala Leu Arg His Val Arg Gln Ile Ala	
370 375 380	
caa gaa aca agt gga gaa gtt cag tat ggt gga ggg cgc caa cct gcg	1200
Gln Glu Thr Ser Gly Glu Val Gln Tyr Gly Gly Gly Arg Gln Pro Ala	
385 390 395 400	
gtt tta aga acc ttc agt caa aga ctc tgt cgg ggt ttc aat gat gct	1248
Val Leu Arg Thr Phe Ser Gln Arg Leu Cys Arg Gly Phe Asn Asp Ala	
405 410 415	
gtt aat ggt ttt gtg gat gat gga tgg tca cca atg ggt agc gat ggt	1296
Val Asn Gly Phe Val Asp Asp Gly Trp Ser Pro Met Gly Ser Asp Gly	
420 425 430	
gca gag gat gtt act gta atg ata aac ttg tcc cct ggg aag ttt ggt	1344
Ala Glu Asp Val Thr Val Met Ile Asn Leu Ser Pro Gly Lys Phe Gly	
435 440 445	
ggg tct cag tac ggt aat tca ttc ctt cca agc ttt ggt agt ggc gtg	1392
Gly Ser Gln Tyr Gly Asn Ser Phe Leu Pro Ser Phe Gly Ser Gly Val	
450 455 460	
ctt tgt gcc aag gca tct atg ttg ctt cag aac gtt cca ccc gct gtg	1440
Leu Cys Ala Lys Ala Ser Met Leu Leu Gln Asn Val Pro Pro Ala Val	
465 470 475 480	
ctg gtt cga ttc ctt aga gaa cac cga tct gaa tgg gct gat tat ggc	1488
Leu Val Arg Phe Leu Arg Glu His Arg Ser Glu Trp Ala Asp Tyr Gly	
485 490 495	
gtg gat gct tat gct gct gca tcg ctc aga gca agt cct ttt gct gtt	1536
Val Asp Ala Tyr Ala Ala Ala Ser Leu Arg Ala Ser Pro Phe Ala Val	
500 505 510	
cct tgt gct aga gct ggg ggg ttc cca agt aac caa gtc att ctt cct	1584
Pro Cys Ala Arg Ala Gly Gly Phe Pro Ser Asn Gln Val Ile Leu Pro	
515 520 525	
ctt gcg cag aca gtt gaa cat gaa gag tca ctt gag gtg gtt aga ctt	1632
Leu Ala Gln Thr Val Glu His Glu Glu Ser Leu Glu Val Val Arg Leu	
530 535 540	
gaa ggt cac gct tac tca ccc gaa gac atg ggt tta gct cgg gat atg	1680
Glu Gly His Ala Tyr Ser Pro Glu Asp Met Gly Leu Ala Arg Asp Met	
545 550 555 560	
tat ttg cta cag ctt tgt agc ggt gtt gat gaa aat gtg gtt gga ggt	1728
Tyr Leu Leu Gln Leu Cys Ser Gly Val Asp Glu Asn Val Val Gly Gly	
565 570 575	
tgt gca cag ctt gta ttt gcc cct atc gat gaa tca ttt gct gat gat	1776
Cys Ala Gln Leu Val Phe Ala Pro Ile Asp Glu Ser Phe Ala Asp Asp	
580 585 590	
gca cct ttg ctt cct tcc ggt ttc cgc atc ata cct ctt gaa cag aaa	1824
Ala Pro Leu Leu Pro Ser Gly Phe Arg Ile Ile Pro Leu Glu Gln Lys	
595 600 605	
tct act ccg aac ggt gca tct gca aac cgt acc ctg gat tta gcc tca	1872
Ser Thr Pro Asn Gly Ala Ser Ala Asn Arg Thr Leu Asp Leu Ala Ser	
610 615 620	
gct tta gaa gga tcc aca cgt caa gct ggt gaa gcc gac cca aat ggc	1920
Ala Leu Glu Gly Ser Thr Arg Gln Ala Gly Glu Ala Asp Pro Asn Gly	

MBI0018 Sequence Listing.ST25

625	630	635	640	
tgt aac ttt agg tcg gta cta acc ata gca ttc cag ttc aca ttt gat				1968
Cys Asn Phe Arg Ser Val Leu Thr Ile Ala Phe Gln Phe Thr Phe Asp	645	650	655	
aac cat tca aga gac agt gtt gct tca atg gca cgt cag tac gtg cga				2016
Asn His Ser Arg Asp Ser Val Ala Ser Met Ala Arg Gln Tyr Val Arg	660	665	670	
agc ata gta gga tcg att cag agg gtt gct cta gcc att gct cct cgt				2064
Ser Ile Val Gly Ser Ile Gln Arg Val Ala Leu Ala Ile Ala Pro Arg	675	680	685	
cct gcc tcc aat atc agt cca ata tct gtt ccc act tcc cct gaa gct				2112
Pro Gly Ser Asn Ile Ser Pro Ile Ser Val Pro Thr Ser Pro Glu Ala	690	695	700	
ctc act ctg gtc cgt tgg atc tcc cgg agt tac agc ctt cac act ggt				2160
Leu Thr Leu Val Arg Trp Ile Ser Arg Ser Tyr Ser Leu His Thr Gly	705	710	715	720
gca gat ctc ttt gga tct gat tct caa acc agt ggt gac acg ttg ctg				2208
Ala Asp Leu Phe Gly Ser Asp Ser Gln Thr Ser Gly Asp Thr Leu Leu	725	730	735	
cat caa ctc tgg aat cac tct gat gca atc ttg tgc tgc tcc ctc aaa				2256
His Gln Leu Trp Asn His Ser Asp Ala Ile Leu Cys Cys Ser Leu Lys	740	745	750	
aca aacgcttcac cggttttcac attcgcaaac caaacgggt tagacatgct				2309
Thr				
ggaaacgact cttgtagccc ttcaagacat aatgctagac aagacccttg acgaacctgg				2369
tcgtaaagct ctttgcctctg agttcccca gatcatgcaa cagggtctatg ctcatctgcc				2429
ggcaggagta tgtgcgtcaa gcatgggaag gatgggtatct tacgagcagg caacgggtgtg				2489
gaaagttctt gaagacgatg aatcaaacca ctgcttagct ttcattgttcg tgaattggtc				2549
gttcgtttga				2559
<210> 42				
<211> 753				
<212> PRT				
<213> Arabidopsis thaliana				
<400> 42				
Met Met Met Val His Ser Met Ser Arg Asp Met Met Asn Arg Glu Ser				
1 5 10 15				
Pro Asp Lys Gly Leu Asp Ser Gly Lys Tyr Val Arg Tyr Thr Pro Glu				
20 25 30				
Gln Val Glu Ala Leu Glu Arg Val Tyr Thr Glu Cys Pro Lys Pro Ser				
35 40 45				
Ser Leu Arg Arg Gln Gln Leu Ile Arg Glu Cys Pro Ile Leu Ser Asn				
50 55 60				
Ile Glu Pro Lys Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg				
65 70 75 80				
Glu Lys Gln Arg Lys Glu Ala Ala Arg Leu Gln Thr Val Asn Arg Lys				
85 90 95				

## MBI0018 Sequence Listing.ST25

Leu Asn Ala Met Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln  
 100 105 110  
 Lys Gln Val Ser Asn Leu Val Tyr Glu Asn Gly His Met Lys His Gln  
 115 120 125  
 Leu His Thr Ala Ser Gly Thr Thr Thr Asp Asn Ser Cys Glu Ser Val  
 130 135 140  
 Val Val Ser Gly Gln Gln His Gln Gln Gln Asn Pro Asn Pro Gln His  
 145 150 155 160  
 Gln Gln Arg Asp Ala Asn Asn Pro Ala Gly Leu Leu Ser Ile Ala Glu  
 165 170 175  
 Glu Ala Leu Ala Glu Phe Leu Ser Lys Ala Thr Gly Thr Ala Val Asp  
 180 185 190  
 Trp Val Gln Met Ile Gly Met Lys Pro Gly Pro Asp Ser Ile Gly Ile  
 195 200 205  
 Val Ala Ile Ser Arg Asn Cys Ser Gly Ile Ala Ala Arg Ala Cys Gly  
 210 215 220  
 Leu Val Ser Leu Glu Pro Met Lys Val Ala Glu Ile Leu Lys Asp Arg  
 225 230 235 240  
 Pro Ser Trp Leu Arg Asp Cys Arg Ser Val Asp Thr Leu Ser Val Ile  
 245 250 255  
 Pro Ala Gly Asn Gly Gly Thr Ile Glu Leu Ile Tyr Thr Gln Met Tyr  
 260 265 270  
 Ala Pro Thr Thr Leu Ala Ala Ala Arg Asp Phe Trp Thr Leu Arg Tyr  
 275 280 285  
 Ser Thr Cys Leu Glu Asp Gly Ser Tyr Val Val Cys Glu Arg Ser Leu  
 290 295 300  
 Thr Ser Ala Thr Gly Gly Pro Thr Gly Pro Pro Ser Ser Asn Phe Val  
 305 310 315 320  
 Arg Ala Glu Met Lys Pro Ser Gly Phe Leu Ile Arg Pro Cys Asp Gly  
 325 330 335  
 Gly Gly Ser Ile Leu His Ile Val Asp His Val Asp Leu Asp Ala Trp  
 340 345 350  
 Ser Val Pro Glu Val Met Arg Pro Leu Tyr Glu Ser Ser Lys Ile Leu  
 355 360 365  
 Ala Gln Lys Met Thr Val Ala Ala Leu Arg His Val Arg Gln Ile Ala  
 370 375 380  
 Gln Glu Thr Ser Gly Glu Val Gln Tyr Gly Gly Gly Arg Gln Pro Ala  
 385 390 395 400

## MBI0018 Sequence Listing.ST25

Val Leu Arg Thr Phe Ser Gln Arg Leu Cys Arg Gly Phe Asn Asp Ala  
 405 410 415  
 Val Asn Gly Phe Val Asp Asp Gly Trp Ser Pro Met Gly Ser Asp Gly  
 420 425 430  
 Ala Glu Asp Val Thr Val Met Ile Asn Leu Ser Pro Gly Lys Phe Gly  
 435 440 445  
 Gly Ser Gln Tyr Gly Asn Ser Phe Leu Pro Ser Phe Gly Ser Gly Val  
 450 455 460  
 Leu Cys Ala Lys Ala Ser Met Leu Leu Gln Asn Val Pro Pro Ala Val  
 465 470 475 480  
 Leu Val Arg Phe Leu Arg Glu His Arg Ser Glu Trp Ala Asp Tyr Gly  
 485 490 495  
 Val Asp Ala Tyr Ala Ala Ala Ser Leu Arg Ala Ser Pro Phe Ala Val  
 500 505 510  
 Pro Cys Ala Arg Ala Gly Gly Phe Pro Ser Asn Gln Val Ile Leu Pro  
 515 520 525  
 Leu Ala Gln Thr Val Glu His Glu Glu Ser Leu Glu Val Val Arg Leu  
 530 535 540  
 Glu Gly His Ala Tyr Ser Pro Glu Asp Met Gly Leu Ala Arg Asp Met  
 545 550 555 560  
 Tyr Leu Leu Gln Leu Cys Ser Gly Val Asp Glu Asn Val Val Gly Gly  
 565 570 575  
 Cys Ala Gln Leu Val Phe Ala Pro Ile Asp Glu Ser Phe Ala Asp Asp  
 580 585 590  
 Ala Pro Leu Leu Pro Ser Gly Phe Arg Ile Ile Pro Leu Glu Gln Lys  
 595 600 605  
 Ser Thr Pro Asn Gly Ala Ser Ala Asn Arg Thr Leu Asp Leu Ala Ser  
 610 615 620  
 Ala Leu Glu Gly Ser Thr Arg Gln Ala Gly Glu Ala Asp Pro Asn Gly  
 625 630 635 640  
 Cys Asn Phe Arg Ser Val Leu Thr Ile Ala Phe Gln Phe Thr Phe Asp  
 645 650 655  
 Asn His Ser Arg Asp Ser Val Ala Ser Met Ala Arg Gln Tyr Val Arg  
 660 665 670  
 Ser Ile Val Gly Ser Ile Gln Arg Val Ala Leu Ala Ile Ala Pro Arg  
 675 680 685  
 Pro Gly Ser Asn Ile Ser Pro Ile Ser Val Pro Thr Ser Pro Glu Ala



MBI0018 Sequence Listing.ST25

690	695	700
Leu Thr Leu Val Arg Trp Ile Ser Arg Ser Tyr Ser Leu His Thr Gly		
705	710	715 720
Ala Asp Leu Phe Gly Ser Asp Ser Gln Thr Ser Gly Asp Thr Leu Leu		
	725	730 735
His Gln Leu Trp Asn His Ser Asp Ala Ile Leu Cys Cys Ser Leu Lys		
	740	745 750

Thr

<210> 43  
 <211> 2526  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)..(2526)  
 <223> G390

<400> 43	
atg atg gct cat cac tcc atg gac gat aga gac tct cct gat aaa gga	48
Met Met Ala His His Ser Met Asp Asp Arg Asp Ser Pro Asp Lys Gly	
1 5 10 15	
ttt gat tcc ggc aag tac gtt aga tac acg ccg gaa caa gtt gaa gct	96
Phe Asp Ser Gly Lys Tyr Val Arg Tyr Thr Pro Glu Gln Val Glu Ala	
20 25 30	
ctt gag aga gtt tat gct gag tgt cct aaa cct agc tct ctg aga aga	144
Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys Pro Ser Ser Leu Arg Arg	
35 40 45	
caa cag ctt att cgt gaa tgt ccc att ctc tgt aac atc gag cct cga	192
Gln Gln Leu Ile Arg Glu Cys Pro Ile Leu Cys Asn Ile Glu Pro Arg	
50 55 60	
cag atc aaa gtt tgg ttc cag aat cgc aga tgt cga gag aag cag agg	240
Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg Glu Lys Gln Arg	
65 70 75 80	
aaa gag tca gct cgt ctt cag aca gtg aac agg aag ctg agt gct atg	288
Lys Glu Ser Ala Arg Leu Gln Thr Val Asn Arg Lys Leu Ser Ala Met	
85 90 95	
aac aag ctt ttg atg gaa gag aat gat cgt ttg cag aag caa gtc tcc	336
Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln Lys Gln Val Ser	
100 105 110	
aac ttg gtt tat gag aat gga ttc atg aaa cat cga atc cac act gct	384
Asn Leu Val Tyr Glu Asn Gly Phe Met Lys His Arg Ile His Thr Ala	
115 120 125	
tct ggg acg acc aca gac aac agc tgt gag tct gtg gtc gtg agt ggt	432
Ser Gly Thr Thr Thr Asp Asn Ser Cys Glu Ser Val Val Val Ser Gly	
130 135 140	
cag caa cgt cag cag caa aac cca aca cat cag cat cct cag cgt gat	480
Gln Gln Arg Gln Gln Gln Asn Pro Thr His Gln His Pro Gln Arg Asp	
145 150 155 160	
gtt aac aac cca gct aat ctt ctc tcg att gcg gag gag acc ttg gcg	528
Val Asn Asn Pro Ala Asn Leu Leu Ser Ile Ala Glu Glu Thr Leu Ala	
165 170 175	

MBI0018 Sequence Listing.ST25																		
gag ttc ctt tgc aag gct aca gga act gct gtc gac tgg gtc cag atg	Glu Phe Leu Cys Lys Ala Thr Gly Thr Ala Val Asp Trp Val Gln Met	180	185	190														576
att ggg atg aag cct ggt ccg gat tct att ggt atc gta gct gtt tca	Ile Gly Met Lys Pro Gly Pro Asp Ser Ile Gly Ile Val Ala Val Ser	195	200	205														624
cgc aac tgc agt gga ata gca gca cgt gcc tgt ggc ctc gtg agt tta	Arg Asn Cys Ser Gly Ile Ala Ala Arg Ala Cys Gly Leu Val Ser Leu	210	215	220														672
gaa ccc atg aag gtc gct gaa atc ctc aaa gat cgt cca tct tgg ttc	Glu Pro Met Lys Val Ala Glu Ile Leu Lys Asp Arg Pro Ser Trp Phe	225	230	235														720
cgt gac tgt cga tgt gtc gag act ctg aat gtt ata ccc act gga aat	Arg Asp Cys Arg Cys Val Glu Thr Leu Asn Val Ile Pro Thr Gly Asn	245	250	255														768
ggt ggt act atc gag ctt gtc aac act cag att tat gct cct aca aca	Gly Gly Thr Ile Glu Leu Val Asn Thr Gln Ile Tyr Ala Pro Thr Thr	260	265	270														816
tta gca gca gct cgt gac ttt tgg acg ctg aga tat agt aca agt cta	Leu Ala Ala Ala Arg Asp Phe Trp Thr Leu Arg Tyr Ser Thr Ser Leu	275	280	285														864
gaa gat gga agc tat gtg gtc tgt gag aga tca ctc act tct gca act	Glu Asp Gly Ser Tyr Val Val Cys Glu Arg Ser Leu Thr Ser Ala Thr	290	295	300														912
ggt ggc ccc aat ggt cca ctt tct tca agc ttc gtg aga gcc aaa atg	Gly Gly Pro Asn Gly Pro Leu Ser Ser Ser Phe Val Arg Ala Lys Met	305	310	315														960
ctg tca agc ggg ttt ctt atc cgt cct tgt gat ggt ggt ggt tcc att	Leu Ser Ser Gly Phe Leu Ile Arg Pro Cys Asp Gly Gly Gly Ser Ile	325	330	335														1008
att cac atc gtt gat cat gtg gac ttg gat gtc tca agt gtt cct gaa	Ile His Ile Val Asp His Val Asp Leu Asp Val Ser Ser Val Pro Glu	340	345	350														1056
gtc ctc agg cct ctt tat gag tct tcc aaa atc ctt gct caa aaa atg	Val Leu Arg Pro Leu Tyr Glu Ser Ser Lys Ile Leu Ala Gln Lys Met	355	360	365														1104
act gtc gct gct ctg aga cat gtg cgc caa att gct caa gag act agt	Thr Val Ala Ala Leu Arg His Val Val Arg Gln Ile Ala Gln Glu Thr Ser	370	375	380														1152
gga gaa gtc cag tat agt ggt gga cgc cag cct gca gtt tta agg act	Gly Glu Val Gln Tyr Ser Gly Gly Arg Gln Pro Ala Val Leu Arg Thr	385	390	395														1200
ttc agc cag aga ctc tgc cgg ggt ttc aat gat gct gta aat ggt ttt	Phe Ser Gln Arg Leu Cys Arg Gly Phe Asn Asp Ala Val Asn Gly Phe	405	410	415														1248
gtc gat gat gga tgg tct cca atg agt agt gat gga gga gag gat att	Val Asp Asp Gly Trp Ser Pro Met Ser Ser Asp Gly Gly Glu Asp Ile	420	425	430														1296
acg atc atg att aac tct tcc tct gct aaa ttt gct ggc tcc caa tac	Thr Ile Met Ile Asn Ser Ser Ser Ala Lys Phe Ala Gly Ser Gln Tyr	435	440	445														1344
ggt agc tca ttt ctt cca agt ttt gga agt ggt gtc ctc tgt gcc aaa	Gly Ser Ser Phe Leu Pro Ser Phe Gly Ser Gly Val Leu Cys Ala Lys	450	455	460														1392
gct tct atg ctg ttg cag aat gtt cca ccc ctt gta ttg att cgg ttc	Ala Ser Met Leu Leu Gln Asn Val Pro Pro Leu Val Leu Ile Arg Phe	465	470	475														1440

## MBI0018 Sequence Listing.ST25

ctg aga gaa cac cga gct gaa tgg gca gac tat ggt gtc gat gcc tat	1488
Leu Arg Glu His Arg Ala Glu Trp Ala Asp Tyr Gly Val Asp Ala Tyr	
485 490 495	
tct gct gca tct ctc aga gca act cca tat gct gtt cca tgc gtc aga	1536
Ser Ala Ala Ser Leu Arg Ala Thr Pro Tyr Ala Val Pro Cys Val Arg	
500 505 510	
acc ggt ggg ttc ccg agt aac caa gtc att ctt cct ctc gca cag aca	1584
Thr Gly Gly Phe Pro Ser Asn Gln Val Ile Leu Pro Leu Ala Gln Thr	
515 520 525	
ctc gaa cat gaa gag ttt ctc gaa gtg gtt aga ctt gga ggt cat gct	1632
Leu Glu His Glu Glu Phe Leu Glu Val Val Arg Leu Gly Gly His Ala	
530 535 540	
tac tca cct gaa gac atg ggc tta tcc cgg gat atg tat tta ctg cag	1680
Tyr Ser Pro Glu Asp Met Gly Leu Ser Arg Asp Met Tyr Leu Leu Gln	
545 550 555 560	
ctt tgt agc ggc gtt gat gaa aat gtg gtt gga ggt tgt gct cag ctt	1728
Leu Cys Ser Gly Val Asp Glu Asn Val Val Gly Gly Cys Ala Gln Leu	
565 570 575	
gtc ttt gcc cca atc gat gaa tca ttt gct gat gat gca cct ttg ctt	1776
Val Phe Ala Pro Ile Asp Glu Ser Phe Ala Asp Asp Ala Pro Leu Leu	
580 585 590	
cct tct ggt ttc cgt gtc ata cca ctc gac caa aaa aca aat ccg aat	1824
Pro Ser Gly Phe Arg Val Ile Pro Leu Asp Gln Lys Thr Asn Pro Asn	
595 600 605	
gat cat caa tct gca agt cga aca cgg gat cta gca tcg tcc cta gat	1872
Asp His Gln Ser Ala Ser Arg Thr Arg Asp Leu Ala Ser Ser Leu Asp	
610 615 620	
ggt tcc acc aaa acc gat tcg gaa aca aac tct aga ttg gtc tta aca	1920
Gly Ser Thr Lys Thr Asp Ser Glu Thr Asn Ser Arg Leu Val Leu Thr	
625 630 635 640	
ata gcc ttc cag ttc acg ttt gat aac cat tcc aga gac aat gtt gct	1968
Ile Ala Phe Gln Phe Thr Phe Asp Asn His Ser Arg Asp Asn Val Ala	
645 650 655	
aca atg gcg aga cag tat gtg agg aac gtt gtt ggt tcg att cag aga	2016
Thr Met Ala Arg Gln Tyr Val Arg Asn Val Val Gly Ser Ile Gln Arg	
660 665 670	
gtg gct cta gcc att acg cct cgt cct ggc tca atg caa ctt ccc act	2064
Val Ala Leu Ala Ile Thr Pro Arg Pro Gly Ser Met Gln Leu Pro Thr	
675 680 685	
tcc cct gaa gct ctc act ctt gtc cgt tgg atc acc cgt agt tac agt	2112
Ser Pro Glu Ala Leu Thr Leu Val Arg Trp Ile Thr Arg Ser Tyr Ser	
690 695 700 705	
att cat aca ggt gca gat ctg ttt gga gct gat tct cag tcc tgt gga	2160
Ile His Thr Gly Ala Asp Leu Phe Gly Ala Asp Ser Gln Ser Cys Gly	
705 710 715 720	
gga gac aca ttg ctt aag caa ctc tgg gac cat agt gat gcc ata ttg	2208
Gly Asp Thr Leu Leu Lys Gln Leu Trp Asp His Ser Asp Ala Ile Leu	
725 730 735	
tgc tgc tcc ctg aaa act aat gcc tca ccg gta ttc aca ttt gca aac	2256
Cys Cys Ser Leu Lys Thr Asn Ala Ser Pro Val Phe Thr Phe Ala Asn	
740 745 750	
caa gct ggt tta gac atg ctt gaa act aca ctt gtg gca ctt cag gat	2304
Gln Ala Gly Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp	
755 760 765	
ata atg ctc gac aaa aca ctt gat gac tct ggt cgt aga gct ctt tgc	2352
Ile Met Leu Asp Lys Thr Leu Asp Asp Ser Gly Arg Arg Ala Leu Cys	

## MBI0018 Sequence Listing.ST25

770	775	780	
tcc gag ttc gcc aag atc atg cag cag gga tat gcg aat ctt ccg gca			2400
Ser Glu Phe Ala Lys Ile Met Gln Gln Gly Tyr Ala Asn Leu Pro Ala			
785	790	795	800
gga ata tgt gtg tcg agc atg ggc aga ccg gtt tcg tat gag caa gcg			2448
Gly Ile Cys Val Ser Ser Met Gly Arg Pro Val Ser Tyr Glu Gln Ala			
	805	810	815
acg gtg tgg aaa gtt gtt gat gac aac gaa tca aac cac tgc ttg gct			2496
Thr Val Trp Lys Val Val Asp Asp Asn Glu Ser Asn His Cys Leu Ala			
	820	825	830
ttt acc ctc gtt agt tgg tcg ttt gtt tga			2526
Phe Thr Leu Val Ser Trp Ser Phe Val			
	835	840	

<210> 44  
 <211> 841  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 44

Met Met Ala His His Ser Met Asp Asp Arg Asp Ser Pro Asp Lys Gly			
1	5	10	15
Phe Asp Ser Gly Lys Tyr Val Arg Tyr Thr Pro Glu Gln Val Glu Ala			
	20	25	30
Leu Glu Arg Val Tyr Ala Glu Cys Pro Lys Pro Ser Ser Leu Arg Arg			
	35	40	45
Gln Gln Leu Ile Arg Glu Cys Pro Ile Leu Cys Asn Ile Glu Pro Arg			
	50	55	60
Gln Ile Lys Val Trp Phe Gln Asn Arg Arg Cys Arg Glu Lys Gln Arg			
	65	70	75
Lys Glu Ser Ala Arg Leu Gln Thr Val Asn Arg Lys Leu Ser Ala Met			
	85	90	95
Asn Lys Leu Leu Met Glu Glu Asn Asp Arg Leu Gln Lys Gln Val Ser			
	100	105	110
Asn Leu Val Tyr Glu Asn Gly Phe Met Lys His Arg Ile His Thr Ala			
	115	120	125
Ser Gly Thr Thr Thr Asp Asn Ser Cys Glu Ser Val Val Val Ser Gly			
	130	135	140
Gln Gln Arg Gln Gln Gln Asn Pro Thr His Gln His Pro Gln Arg Asp			
	145	150	155
Val Asn Asn Pro Ala Asn Leu Leu Ser Ile Ala Glu Glu Thr Leu Ala			
	165	170	175
Glu Phe Leu Cys Lys Ala Thr Gly Thr Ala Val Asp Trp Val Gln Met			
	180	185	190
Ile Gly Met Lys Pro Gly Pro Asp Ser Ile Gly Ile Val Ala Val Ser			

	195					200					205				
Arg	Asn 210	Cys	Ser	Gly	Ile	Ala 215	Ala	Arg	Ala	Cys	Gly 220	Leu	Val	Ser	Leu
Glu 225	Pro	Met	Lys	Val	Ala 230	Glu	Ile	Leu	Lys	Asp 235	Arg	Pro	Ser	Trp	Phe 240
Arg	Asp	Cys	Arg	Cys 245	Val	Glu	Thr	Leu	Asn 250	Val	Ile	Pro	Thr	Gly 255	Asn
Gly	Gly	Thr	Ile 260	Glu	Leu	Val	Asn 265	Thr	Gln	Ile	Tyr	Ala	Pro 270	Thr	Thr
Leu	Ala	Ala 275	Ala	Arg	Asp	Phe	Trp 280	Thr	Leu	Arg	Tyr	Ser 285	Thr	Ser	Leu
Glu	Asp 290	Gly	Ser	Tyr	Val	Val 295	Cys	Glu	Arg	Ser	Leu 300	Thr	Ser	Ala	Thr
Gly 305	Gly	Pro	Asn	Gly	Pro 310	Leu	Ser	Ser	Ser	Phe 315	Val	Arg	Ala	Lys	Met 320
Leu	Ser	Ser	Gly	Phe 325	Leu	Ile	Arg	Pro	Cys 330	Asp	Gly	Gly	Gly	Ser 335	Ile
Ile	His	Ile	Val 340	Asp	His	Val	Asp	Leu 345	Asp	Val	Ser	Ser	Val 350	Pro	Glu
Val	Leu	Arg 355	Pro	Leu	Tyr	Glu	Ser 360	Ser	Lys	Ile	Leu	Ala 365	Gln	Lys	Met
Thr	Val 370	Ala	Ala	Leu	Arg	His 375	Val	Arg	Gln	Ile	Ala 380	Gln	Glu	Thr	Ser
Gly 385	Glu	Val	Gln	Tyr	Ser 390	Gly	Gly	Arg	Gln	Pro 395	Ala	Val	Leu	Arg	Thr 400
Phe	Ser	Gln	Arg	Leu 405	Cys	Arg	Gly	Phe	Asn 410	Asp	Ala	Val	Asn	Gly 415	Phe
Val	Asp	Asp	Gly 420	Trp	Ser	Pro	Met	Ser 425	Ser	Asp	Gly	Gly	Glu 430	Asp	Ile
Thr	Ile 435	Met	Ile	Asn	Ser	Ser	Ser 440	Ala	Lys	Phe	Ala	Gly 445	Ser	Gln	Tyr
Gly	Ser 450	Ser	Phe	Leu	Pro	Ser	Phe 455	Gly	Ser	Gly	Val 460	Leu	Cys	Ala	Lys
Ala 465	Ser	Met	Leu	Leu	Gln 470	Asn	Val	Pro	Pro	Leu 475	Val	Leu	Ile	Arg	Phe 480
Leu	Arg	Glu	His	Arg 485	Ala	Glu	Trp	Ala	Asp 490	Tyr	Gly	Val	Asp	Ala 495	Tyr

## MBI0018 Sequence Listing.ST25

Ser Ala Ala Ser Leu Arg Ala Thr Pro Tyr Ala Val Pro Cys Val Arg  
 500 505 510  
 Thr Gly Gly Phe Pro Ser Asn Gln Val Ile Leu Pro Leu Ala Gln Thr  
 515 520 525  
 Leu Glu His Glu Glu Phe Leu Glu Val Val Arg Leu Gly Gly His Ala  
 530 535 540  
 Tyr Ser Pro Glu Asp Met Gly Leu Ser Arg Asp Met Tyr Leu Leu Gln  
 545 550 555 560  
 Leu Cys Ser Gly Val Asp Glu Asn Val Val Gly Gly Cys Ala Gln Leu  
 565 570 575  
 Val Phe Ala Pro Ile Asp Glu Ser Phe Ala Asp Asp Ala Pro Leu Leu  
 580 585 590  
 Pro Ser Gly Phe Arg Val Ile Pro Leu Asp Gln Lys Thr Asn Pro Asn  
 595 600 605  
 Asp His Gln Ser Ala Ser Arg Thr Arg Asp Leu Ala Ser Ser Leu Asp  
 610 615 620  
 Gly Ser Thr Lys Thr Asp Ser Glu Thr Asn Ser Arg Leu Val Leu Thr  
 625 630 635 640  
 Ile Ala Phe Gln Phe Thr Phe Asp Asn His Ser Arg Asp Asn Val Ala  
 645 650 655  
 Thr Met Ala Arg Gln Tyr Val Arg Asn Val Val Gly Ser Ile Gln Arg  
 660 665 670  
 Val Ala Leu Ala Ile Thr Pro Arg Pro Gly Ser Met Gln Leu Pro Thr  
 675 680 685  
 Ser Pro Glu Ala Leu Thr Leu Val Arg Trp Ile Thr Arg Ser Tyr Ser  
 690 695 700  
 Ile His Thr Gly Ala Asp Leu Phe Gly Ala Asp Ser Gln Ser Cys Gly  
 705 710 715 720  
 Gly Asp Thr Leu Leu Lys Gln Leu Trp Asp His Ser Asp Ala Ile Leu  
 725 730 735  
 Cys Cys Ser Leu Lys Thr Asn Ala Ser Pro Val Phe Thr Phe Ala Asn  
 740 745 750  
 Gln Ala Gly Leu Asp Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp  
 755 760 765  
 Ile Met Leu Asp Lys Thr Leu Asp Asp Ser Gly Arg Arg Ala Leu Cys  
 770 775 780  
 Ser Glu Phe Ala Lys Ile Met Gln Gln Gly Tyr Ala Asn Leu Pro Ala  
 785 790 795 800

## MBI0018 Sequence Listing.ST25

Gly Ile Cys Val Ser Ser Met Gly Arg Pro Val Ser Tyr Glu Gln Ala  
805 810 815

Thr Val Trp Lys Val Val Asp Asp Asn Glu Ser Asn His Cys Leu Ala  
820 825 830

Phe Thr Leu Val Ser Trp Ser Phe Val  
835 840

<210> 45  
<211> 2511  
<212> DNA  
<213> Arabidopsis thaliana

<220>  
<221> CDS  
<222> (1)..(2511)  
<223> G1548

<400> 45  
atg gca atg tct tgc aag gat ggt aag ttg gga tgt ttg gat aat ggg 48  
Met Ala Met Ser Cys Lys Asp Gly Lys Leu Gly Cys Leu Asp Asn Gly  
1 5 10 15  
aag tat gtg agg tat aca cct gaa caa gtt gaa gca ctt gag agg ctt 96  
Lys Tyr Val Arg Tyr Thr Pro Glu Gln Val Glu Ala Leu Glu Arg Leu  
20 25 30  
tat cat gac tgt cct aaa ccg agt tct att cgc cgt cag cag ttg atc 144  
Tyr His Asp Cys Pro Lys Pro Ser Ser Ile Arg Arg Gln Gln Leu Ile  
35 40 45  
aga gag tgt cct att ctc tct aac att gag cct aaa cag atc aaa gtg 192  
Arg Glu Cys Pro Ile Leu Ser Asn Ile Glu Pro Lys Gln Ile Lys Val  
50 55 60  
tgg ttt cag aac cga aga tgt aga gag aaa caa agg aaa gag gct tca 240  
Trp Phe Gln Asn Arg Arg Cys Arg Glu Lys Gln Arg Lys Glu Ala Ser  
65 70 75 80  
cgg ctt caa gct gtg aat cgg aag ttg acg gca atg aac aag ctc ttg 288  
Arg Leu Gln Ala Val Asn Arg Lys Leu Thr Ala Met Asn Lys Leu Leu  
85 90 95  
atg gag gag aat gac agg ttg cag aag caa gtg tca cag ctg gtc cat 336  
Met Glu Glu Asn Asp Arg Leu Gln Lys Gln Val Ser Gln Leu Val His  
100 105 110  
gaa aac agc tac ttc cgt caa cat act cca aat cct tca ctc cca gct 384  
Glu Asn Ser Tyr Phe Arg Gln His Thr Pro Asn Pro Ser Leu Pro Ala  
115 120 125  
aaa gac aca agc tgt gaa tcg gtg gtg acg agt ggt cag cac caa ttg 432  
Lys Asp Thr Ser Cys Glu Ser Val Val Thr Ser Gly Gln His Gln Leu  
130 135 140  
gca tct caa aat cct cag aga gat gct agt cct gca gga ctt ttg tcc 480  
Ala Ser Gln Asn Pro Gln Arg Asp Ala Ser Pro Ala Gly Leu Leu Ser  
145 150 155 160  
att gca gaa gaa act tta gca gag ttt ctt tca aag gca act gga acc 528  
Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala Thr Gly Thr  
165 170 175  
gct gtt gag tgg gtt cag atg cct gga atg aag cct ggt ccg gat tcc 576  
Ala Val Glu Trp Val Gln Met Pro Gly Met Lys Pro Gly Pro Asp Ser  
180 185 190  
att gga atc atc gct att tct cat ggt tgc act ggt gtg gca gca cgc 624  
Ile Gly Ile Ile Ala Ile Ser His Gly Cys Thr Gly Val Ala Ala Arg

MBI0018 Sequence Listing.ST25																
195					200					205						
gcc Ala	tgt Cys 210	ggc Gly	cta Leu	gtg Val	ggc Gly	ctt Leu 215	gag Glu	cct Pro	aca Thr	agg Arg	gtt Val 220	gca Ala	gag Glu	att Ile	gtc Val	672
aag Lys 225	gat Asp	cgt Arg	cct Pro	tcg Ser	tgg Trp 230	ttc Phe	cgc Arg	gaa Glu	tgt Cys	cga Arg 235	gct Ala	gtt Val	gaa Glu	gtt Val	atg Met 240	720
aac Asn	gtg Val	ttg Leu	cca Pro	act Thr 245	gcc Ala	aat Asn	ggc Gly	gga Gly	acc Thr 250	gtt Val	gag Glu	ctg Leu	ctt Leu	tat Tyr 255	atg Met	768
cag Gln	ctc Leu	tat Tyr	gca Ala 260	cca Pro	act Thr	aca Thr	ttg Leu	gcc Ala 265	cca Pro	cca Pro	cgc Arg	gat Asp	ttc Phe 270	tgg Trp	ctg Leu	816
tta Leu	cgt Arg	tac Tyr 275	acc Thr	tct Ser	gtt Val	tta Leu	gaa Glu 280	gat Asp	ggc Gly	agc Ser	ctt Leu	gtg Val 285	gtg Val	tgc Cys	gag Glu	864
aga Arg	tct Ser 290	ctt Leu	aag Lys	agc Ser	act Thr	caa Gln 295	aat Asn	ggc Gly	cct Pro	agt Ser	atg Met 300	cca Pro	ctg Leu	gtt Val	cag Gln	912
aat Asn 305	ttt Phe	gtg Val	aga Arg	gca Ala	gag Glu 310	atg Met	ctt Leu	tcc Ser	agt Ser	ggg Gly 315	tac Tyr	ttg Leu	ata Ile	cgg Arg	cct Pro 320	960
tgt Cys	gat Asp	ggc Gly	ggc Gly	ggc Gly 325	tca Ser	atc Ile	ata Ile	cac His	ata Ile 330	gtg Val	gat Asp	cat His	atg Met	gat Asp 335	ttg Leu	1008
gag Glu	gct Ala	tgt Cys	agc Ser 340	gtg Val	cct Pro	gag Glu	gtc Val	ttg Leu 345	cgc Arg	ccg Pro	ctc Leu	tat Tyr	gag Glu 350	tca Ser	ccc Pro	1056
aaa Lys	gta Val	ctt Leu 355	gca Ala	cag Gln	aag Lys	aca Thr	aca Thr 360	atg Met	gcg Ala	gca Ala	ctg Leu	cgt Arg 365	cag Gln	ctc Leu	aag Lys	1104
caa Gln	ata Ile 370	gct Ala	cag Gln	gag Glu	gtt Val	act Thr 375	cag Gln	act Thr	aat Asn	agt Ser	agt Ser 380	gtt Val	aat Asn	ggg Gly	tgg Trp	1152
gga Gly 385	cgg Arg	cgt Arg	cct Pro	gct Ala	gcc Ala 390	tta Leu	aga Arg	gct Ala	ctc Leu	agc Ser 395	cag Gln	agg Arg	cta Leu	agc Ser	aga Arg 400	1200
ggc Gly	ttc Phe	aat Asn	gaa Glu 405	gct Ala	gta Val	aat Asn	ggc Gly	ttc Phe	act Thr 410	gat Asp	gaa Glu	gga Gly	tgg Trp	tca Ser 415	gtg Val	1248
ata Ile	gga Gly	gat Asp	agc Ser 420	atg Met	gat Asp	gat Asp	gtc Val	aca Thr 425	atc Ile	act Thr	gta Val	aac Asn	tct Ser 430	tct Ser	cca Pro	1296
gac Asp	aag Lys	cta Leu 435	atg Met	ggc Gly	cta Leu	aat Asn	ctt Leu 440	aca Thr	ttt Phe	gcc Ala	aat Asn	ggc Gly 445	ttt Phe	gct Ala	cct Pro	1344
gta Val	agc Ser 450	aat Asn	gtt Val	gtt Val	tta Leu	tgc Cys 455	gca Ala	aaa Lys	gca Ala	tca Ser	atg Met 460	ctt Leu	tta Leu	cag Gln	aat Asn	1392
gtt Val 465	cct Pro	ccg Pro	gcg Ala	atc Ile	ctg Leu 470	ctt Leu	cgg Arg	ttt Phe	ctg Leu	agg Arg 475	gag Glu	cat His	agg Arg	tca Ser	gaa Glu 480	1440
tgg Trp	gct Ala	gac Asp	aac Asn	aac Asn 485	att Ile	gat Asp	gcg Ala	tat Tyr	cta Leu 490	gca Ala	gca Ala	gca Ala	gtt Val	aaa Lys 495	gta Val	1488
qqq	cct	tgt	aqt	qcc	cqa	qtt	qga	qga	ttt	qga	ggg	cag	gtt	ata	ctt	1536



MBI0018 Sequence Listing.ST25

Gly	Pro	Cys	Ser	Ala	Arg	Val	Gly	Gly	Phe	Gly	Gly	Gln	Val	Ile	Leu	
			500					505							510	
cca	ctt	gct	cat	act	att	gag	cat	gaa	gag	ttt	atg	gaa	gtc	atc	aaa	1584
Pro	Leu	Ala	His	Thr	Ile	Glu	His	Glu	Glu	Phe	Met	Glu	Val	Ile	Lys	
		515						520				525				
ttg	gaa	ggg	ctt	ggg	cat	tcc	cct	gaa	gat	gca	atc	gtt	cca	aga	gat	1632
Leu	Glu	Gly	Leu	Gly	His	Ser	Pro	Glu	Asp	Ala	Ile	Val	Pro	Arg	Asp	
	530					535					540					
atc	ttc	ctt	ctt	caa	ctt	tgt	agc	gga	atg	gat	gaa	aat	gct	gta	gga	1680
Ile	Phe	Leu	Leu	Gln	Leu	Cys	Ser	Gly	Met	Asp	Glu	Asn	Ala	Val	Gly	
					550					555					560	
acc	tgt	gcg	gaa	ctt	ata	ttt	gct	cca	atc	gat	gct	tcg	ttt	gcg	gat	1728
Thr	Cys	Ala	Glu	Leu	Ile	Phe	Ala	Pro	Ile	Asp	Ala	Ser	Phe	Ala	Asp	
				565					570					575		
gat	gca	cct	ctg	ctt	cct	tct	ggg	ttt	cgt	att	atc	cct	ctt	gat	tcc	1776
Asp	Ala	Pro	Leu	Leu	Pro	Ser	Gly	Phe	Arg	Ile	Ile	Pro	Leu	Asp	Ser	
			580					585					590			
gca	aag	gaa	gta	tct	agc	cca	aac	cga	acc	ttg	gat	ctt	gct	tcg	gca	1824
Ala	Lys	Glu	Val	Ser	Ser	Pro	Asn	Arg	Thr	Leu	Asp	Leu	Ala	Ser	Ala	
		595					600					605				
ctg	gaa	att	ggg	tca	gct	gga	aca	aaa	gcc	tca	act	gat	caa	tca	gga	1872
Leu	Glu	Ile	Gly	Ser	Ala	Gly	Thr	Lys	Ala	Ser	Thr	Asp	Gln	Ser	Gly	
	610					615					620					
aac	tcc	aca	tgt	gca	aga	tct	gtg	atg	aca	ata	gca	ttt	gag	ttt	ggg	1920
Asn	Ser	Thr	Cys	Ala	Arg	Ser	Val	Met	Thr	Ile	Ala	Phe	Glu	Phe	Gly	
	625				630					635					640	
atc	gag	agc	cat	atg	caa	gaa	cat	gta	gca	tcc	atg	gct	agg	cag	tat	1968
Ile	Glu	Ser	His	Met	Gln	Glu	His	Val	Ala	Ser	Met	Ala	Arg	Gln	Tyr	
				645					650					655		
gtt	cga	ggg	atc	ata	tca	tcg	gtg	cag	aga	gta	gca	ttg	gct	ctt	tct	2016
Val	Arg	Gly	Ile	Ile	Ser	Ser	Val	Gln	Arg	Val	Ala	Leu	Ala	Leu	Ser	
			660					665					670			
cct	tct	cat	atc	agc	tca	caa	gtt	ggg	cta	cgc	act	cct	ttg	ggg	act	2064
Pro	Ser	His	Ile	Ser	Ser	Gln	Val	Gly	Leu	Arg	Thr	Pro	Leu	Gly	Thr	
		675					680					685				
cct	gaa	gcc	caa	aca	ctt	gct	cgt	tgg	att	tgc	cag	agt	tac	agg	ggc	2112
Pro	Glu	Ala	Gln	Thr	Leu	Ala	Arg	Trp	Ile	Cys	Gln	Ser	Tyr	Arg	Gly	
	690					695				700						
tac	atg	ggg	gtt	gag	cta	ctt	aaa	tca	aac	agt	gac	ggc	aat	gaa	tct	2160
Tyr	Met	Gly	Val	Glu	Leu	Leu	Lys	Ser	Asn	Ser	Asp	Gly	Asn	Glu	Ser	
	705				710					715					720	
att	ctt	aag	aat	ctt	tgg	cat	cac	act	gat	gct	ata	atc	tgc	tgc	tca	2208
Ile	Leu	Lys	Asn	Leu	Trp	His	His	Thr	Asp	Ala	Ile	Ile	Cys	Cys	Ser	
			725						730				735			
atg	aag	gcc	ttg	ccc	gtc	ttc	aca	ttt	gca	aac	cag	gcg	gga	ctt	gac	2256
Met	Lys	Ala	Leu	Pro	Val	Phe	Thr	Phe	Ala	Asn	Gln	Ala	Gly	Leu	Asp	
			740					745					750			
atg	ctg	gag	act	aca	tta	gtt	gct	ctt	caa	gac	atc	tct	tta	gag	aag	2304
Met	Leu	Glu	Thr	Thr	Leu	Val	Ala	Leu	Gln	Asp	Ile	Ser	Leu	Glu	Lys	
		755					760					765				
ata	ttt	gat	gac	aat	gga	aga	aag	act	ctt	tgc	tct	gag	ttc	cca	cag	2352
Ile	Phe	Asp	Asp	Asn	Gly	Arg	Lys	Thr	Leu	Cys	Ser	Glu	Phe	Pro	Gln	
	770					775					780					
atc	atg	caa	cag	ggc	ttc	gcg	tgc	ctt	caa	ggc	ggg	ata	tgt	ctc	tca	2400
Ile	Met	Gln	Gln	Gly	Phe	Ala	Cys	Leu	Gln	Gly	Gly	Ile	Cys	Leu	Ser	
	785				790					795					800	

MBI0018 Sequence Listing.ST25

agc atg ggg aga cca gtt tcg tat gag aga gca gtt gct tgg aaa gta	2448
Ser Met Gly Arg Pro Val Ser Tyr Glu Arg Ala Val Ala Trp Lys Val	
805 810 815	
ctc aat gaa gaa gaa aat gct cat tgc atc tgc ttt gtg ttc atc aat	2496
Leu Asn Glu Glu Glu Asn Ala His Cys Ile Cys Phe Val Phe Ile Asn	
820 825 830	
tgg tcc ttt gtg tga	2511
Trp Ser Phe Val	
835	
<210> 46	
<211> 836	
<212> PRT	
<213> Arabidopsis thaliana	
<400> 46	
Met Ala Met Ser Cys Lys Asp Gly Lys Leu Gly Cys Leu Asp Asn Gly	
1 5 10 15	
Lys Tyr Val Arg Tyr Thr Pro Glu Gln Val Glu Ala Leu Glu Arg Leu	
20 25 30	
Tyr His Asp Cys Pro Lys Pro Ser Ser Ile Arg Arg Gln Gln Leu Ile	
35 40 45	
Arg Glu Cys Pro Ile Leu Ser Asn Ile Glu Pro Lys Gln Ile Lys Val	
50 55 60	
Trp Phe Gln Asn Arg Arg Cys Arg Glu Lys Gln Arg Lys Glu Ala Ser	
65 70 75 80	
Arg Leu Gln Ala Val Asn Arg Lys Leu Thr Ala Met Asn Lys Leu Leu	
85 90 95	
Met Glu Glu Asn Asp Arg Leu Gln Lys Gln Val Ser Gln Leu Val His	
100 105 110	
Glu Asn Ser Tyr Phe Arg Gln His Thr Pro Asn Pro Ser Leu Pro Ala	
115 120 125	
Lys Asp Thr Ser Cys Glu Ser Val Val Thr Ser Gly Gln His Gln Leu	
130 135 140	
Ala Ser Gln Asn Pro Gln Arg Asp Ala Ser Pro Ala Gly Leu Leu Ser	
145 150 155 160	
Ile Ala Glu Glu Thr Leu Ala Glu Phe Leu Ser Lys Ala Thr Gly Thr	
165 170 175	
Ala Val Glu Trp Val Gln Met Pro Gly Met Lys Pro Gly Pro Asp Ser	
180 185 190	
Ile Gly Ile Ile Ala Ile Ser His Gly Cys Thr Gly Val Ala Ala Arg	
195 200 205	
Ala Cys Gly Leu Val Gly Leu Glu Pro Thr Arg Val Ala Glu Ile Val	
210 215 220	

## MBI0018 Sequence Listing.ST25

Lys Asp Arg Pro Ser Trp Phe Arg Glu Cys Arg Ala Val Glu Val Met  
 225 230 235 240  
 Asn Val Leu Pro Thr Ala Asn Gly Gly Thr Val Glu Leu Leu Tyr Met  
 245 250 255  
 Gln Leu Tyr Ala Pro Thr Thr Leu Ala Pro Pro Arg Asp Phe Trp Leu  
 260 265 270  
 Leu Arg Tyr Thr Ser Val Leu Glu Asp Gly Ser Leu Val Val Cys Glu  
 275 280 285  
 Arg Ser Leu Lys Ser Thr Gln Asn Gly Pro Ser Met Pro Leu Val Gln  
 290 295 300  
 Asn Phe Val Arg Ala Glu Met Leu Ser Ser Gly Tyr Leu Ile Arg Pro  
 305 310 315 320  
 Cys Asp Gly Gly Gly Ser Ile Ile His Ile Val Asp His Met Asp Leu  
 325 330 335  
 Glu Ala Cys Ser Val Pro Glu Val Leu Arg Pro Leu Tyr Glu Ser Pro  
 340 345 350  
 Lys Val Leu Ala Gln Lys Thr Thr Met Ala Ala Leu Arg Gln Leu Lys  
 355 360 365  
 Gln Ile Ala Gln Glu Val Thr Gln Thr Asn Ser Ser Val Asn Gly Trp  
 370 375 380  
 Gly Arg Arg Pro Ala Ala Leu Arg Ala Leu Ser Gln Arg Leu Ser Arg  
 385 390 395 400  
 Gly Phe Asn Glu Ala Val Asn Gly Phe Thr Asp Glu Gly Trp Ser Val  
 405 410 415  
 Ile Gly Asp Ser Met Asp Asp Val Thr Ile Thr Val Asn Ser Ser Pro  
 420 425 430  
 Asp Lys Leu Met Gly Leu Asn Leu Thr Phe Ala Asn Gly Phe Ala Pro  
 435 440 445  
 Val Ser Asn Val Val Leu Cys Ala Lys Ala Ser Met Leu Leu Gln Asn  
 450 455 460  
 Val Pro Pro Ala Ile Leu Leu Arg Phe Leu Arg Glu His Arg Ser Glu  
 465 470 475 480  
 Trp Ala Asp Asn Asn Ile Asp Ala Tyr Leu Ala Ala Ala Val Lys Val  
 485 490 495  
 Gly Pro Cys Ser Ala Arg Val Gly Gly Phe Gly Gly Gln Val Ile Leu  
 500 505 510  
 Pro Leu Ala His Thr Ile Glu His Glu Glu Phe Met Glu Val Ile Lys  
 515 520 525

## MBI0018 Sequence Listing.ST25

Leu Glu Gly Leu Gly His Ser Pro Glu Asp Ala Ile Val Pro Arg Asp  
 530 535 540  
 Ile Phe Leu Leu Gln Leu Cys Ser Gly Met Asp Glu Asn Ala Val Gly  
 545 550 555 560  
 Thr Cys Ala Glu Leu Ile Phe Ala Pro Ile Asp Ala Ser Phe Ala Asp  
 565 570 575  
 Asp Ala Pro Leu Leu Pro Ser Gly Phe Arg Ile Ile Pro Leu Asp Ser  
 580 585 590  
 Ala Lys Glu Val Ser Ser Pro Asn Arg Thr Leu Asp Leu Ala Ser Ala  
 595 600 605  
 Leu Glu Ile Gly Ser Ala Gly Thr Lys Ala Ser Thr Asp Gln Ser Gly  
 610 615 620  
 Asn Ser Thr Cys Ala Arg Ser Val Met Thr Ile Ala Phe Glu Phe Gly  
 625 630 635 640  
 Ile Glu Ser His Met Gln Glu His Val Ala Ser Met Ala Arg Gln Tyr  
 645 650 655  
 Val Arg Gly Ile Ile Ser Ser Val Gln Arg Val Ala Leu Ala Leu Ser  
 660 665 670  
 Pro Ser His Ile Ser Ser Gln Val Gly Leu Arg Thr Pro Leu Gly Thr  
 675 680 685  
 Pro Glu Ala Gln Thr Leu Ala Arg Trp Ile Cys Gln Ser Tyr Arg Gly  
 690 695 700  
 Tyr Met Gly Val Glu Leu Leu Lys Ser Asn Ser Asp Gly Asn Glu Ser  
 705 710 715 720  
 Ile Leu Lys Asn Leu Trp His His Thr Asp Ala Ile Ile Cys Cys Ser  
 725 730 735  
 Met Lys Ala Leu Pro Val Phe Thr Phe Ala Asn Gln Ala Gly Leu Asp  
 740 745 750  
 Met Leu Glu Thr Thr Leu Val Ala Leu Gln Asp Ile Ser Leu Glu Lys  
 755 760 765  
 Ile Phe Asp Asp Asn Gly Arg Lys Thr Leu Cys Ser Glu Phe Pro Gln  
 770 775 780  
 Ile Met Gln Gln Gly Phe Ala Cys Leu Gln Gly Gly Ile Cys Leu Ser  
 785 790 795 800  
 Ser Met Gly Arg Pro Val Ser Tyr Glu Arg Ala Val Ala Trp Lys Val  
 805 810 815  
 Leu Asn Glu Glu Glu Asn Ala His Cys Ile Cys Phe Val Phe Ile Asn

820 MBI0018 Sequence Listing.ST25  
825 830

Trp Ser Phe Val  
835

# INTERNATIONAL SEARCH REPORT

national application No.

PCT/US00/31325

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C07H 21/04; C12N 5/10, 15/29, 15/63, 15/82  
US CL. : 435/320.1, 419, 440, 468; 536/23.1, 23.6; 800/278, 290

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 419, 440, 468; 536/23.1, 23.6; 800/278, 290

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Continuation Sheet

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database Genbank on NCBI, US National Library of Medicine (Bethesda, MD, USA). No. AJ005196, BUCHHEITZ, G. et al. 'Nuclear-localized receiver-like proteins are differentially expressed in Arabidopsis thaliana'. September 4, 1998.	4,6,9,10
Y	SAKAI, H. et al. Two-component response regulators from Arabidopsis thaliana contain a putative DNA-binding motif. Plant Cell Physiology 1998, Vol 39 No. 11, pages 1232-1239, see entire document.	1-3,5,7,8,9,13,27-27
X	GILOVER, B.J. et al. Development of several epidermal cell types can be specified by the same MYB-related plant transcription factor. Development 1998, Vol 125, pages 3497-3508, see entire document.	4,6
Y	MARTIN, C. et al. MYB transcription factors in plants. Trends in Genetics, February 1997, Vol 13, No 2, pages 67-73, see entire document.	1-10, 13, 25-27
A		1-10, 13, 25-27

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"I" earlier application or patent published on or after the international filing date

"I" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

04 April 2001 (04.04.2001)

Date of completion of the international search report

03 MAY 2001

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

David H Kruse

Telephone No. 703-308-0196

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/31325

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claim Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claim Nos.: 14 and 23  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:  
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.: 1-10,13,25-27 and SEQ ID NO: 1,2,29&30
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

### Remark on Protest

☐  
☒

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/31325

**BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING** This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I-XXIII, claim(s) 1-10, 13, 14 and 25-27, drawn to a transgenic plant having modified structure and development characteristics, polynucleotides and vectors for producing said transgenic plant and a method of making said transgenic plant. Applicant must elect one pair of sequences (one nucleic acid and the corresponding amino acid translation) to be examined, *i.e.* SEQ ID NO: 1 and 2 in Group I, SEQ ID NO: 3 and 4 in Group II, SEQ ID NO: 5 and 6 in Group III, etc.

Group XXIV, claim(s) 11 and 12, drawn to an isolated or recombinant polypeptide.

Group XXV, claim(s) 15-17, drawn to a method of identifying a factor that is modulated by or interacts with a polypeptide.

Group XXVI, claim(s) 18, drawn to a method of identifying a molecule that modulates activity or expression of a polynucleotide or polypeptide.

Group XXVII, claim(s) 19 and 20, drawn to an integrated data system.

Group XXVIII, claim(s) 21-24, drawn to a method of identifying a polynucleotide or polypeptide sequence homologue.

The inventions listed as Groups I-XXVIII do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The inventions listed as Groups I-XXVIII do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Groups I-XXIII are drawn to a transgenic plant and a method of producing said plant with a nucleic acid sequence encoding a wide variety of transcription factors. Group XXIV is drawn to a wide variety of isolated or recombinant polypeptides having transcriptional factor activity. The methods of Groups I-XXIII differ from each other in that they are directed to a plant transformation method and transgenic plant with a structurally and functionally distinct nucleic acid sequence which encodes a structurally and functionally distinct amino acid sequence. In addition, Groups XXV, XXVI, XXVII and XXVIII are different methods from any of Groups I-XXIII in that they have different method steps and different end products, and Group XXVII requires a computer system. Thus, there is no single special technical feature, which links the inventions of Groups I-XXVIII under PCT Rule 13.2.

**Continuation of B. FIELDS SEARCHED Item 3:** EAST (USPAT); STN (AGRICOLA, BIOSIS, CAPLUS, EMBASE); Sequence Search SEQ ID NO: 1, 2; 29 and 30; NCBI/Genbank.